Pricing in Linear-Quadratic Dynamic Games

Lillian J. Ratliff, Samuel Coogan, Daniel Calderone, and S. Shankar Sastry

Abstract— **We investigate the use of pricing mechanisms as a means to achieve a desired feedback control strategy among selfish agents. We study a hierarchical linear-quadratic game with many dynamically coupled Nash followers and an uncoupled leader. The leader influences the game by choosing the quadratic dependence on control actions for each follower's cost function. We show that determining whether the leader can establish the desired feedback control as a Nash equilibrium among the followers is a convex feasibility problem for the continuous-time infinite horizon, discrete-time infinite horizon, and discrete-time finite horizon settings, and we present several extensions to this main result. In particular, we discuss methods for ensuring that the total cost incurred due to the leader's pricing is as close as possible to a specified nominal cost, as well as methods for minimizing the explicit dependence of a player's cost on other players' control inputs. Finally, we apply the proposed method to the problem of ensuring the security of a multi-network and to the problem of pricing for controlled diffusion in a general network.**

I. INTRODUCTION

In many engineering problems it is common for there to be multiple decision making agents with competing interests. It is important to accurately model these systems and develop resilient control strategies that account for the interests of all the participating agents while meeting the organizational objective which may represent social welfare or the common good. In systems where agents interact in a noncooperative manner and there are socioeconomic considerations, game theory can be a powerful tool for modeling agent interaction and designing mechanisms to coordinate agents. From a noncooperative, distributed control point of view, pricing mechanisms allow for the leader to close the gap between the centralized cost, which is equivalent to the cost incurred when all agents play the socially optimal strategy, and the decentralized cost, which is equivalent to the cost incurred when each follower plays a purely selfish strategy.

There has been considerable literature proposing the use of mechanisms as a means to achieve a socially optimal solution [1]–[4]. The idea of designing local utility functions to achieve a socially optimal solution to a multi-agent game is not a new idea [5], [6]. While the problem of designing utilities is considered, it is applied to the specific problem of distribution of welfare. In some cases, the problem is only considered in the static or finite game context. In [7], the authors consider utility design to achieve a global objective while obeying coupling constraints. The authors propose augmenting the system to form a state-based game when such utilities do not exist.

We propose the use of pricing schemes as a means to encourage cooperation among selfish agents in order to achieve a societally optimal solution. The particular form of the problem that we study is a hierarchical, linear-quadratic dynamic game with one leader and many Nash followers. The class of pricing mechanisms we consider are quadratic in the choice variables of the followers. The main difference between the research proposed in this paper and the existing literature is that we make a connection between mechanism design and inverse optimal control for hierarchical, linear-quadratic dynamic games and we formulate the problem of designing pricing mechanisms for linear-quadratic dynamic games in discrete and continuous time as a convex feasibility problem. We apply the developed theory to the problems of network security and controlled diffusion in a multi-agent network.

The pricing mechanisms ensure the followers act according to a desired Nash equilibrium strategy. Since we consider a dynamic setting we do not lose the temporal information of the game. Further, we provide a pricing scheme that is guaranteed to enforce control strategies that are functions of the system state which evolves according to a dynamical system. In addition, we propose objectives which can be optimized in order to select pricing mechanisms with desirable features. In the network security application, the pricing mechanisms are introduced in order to induce agents to invest in security. In the controlled diffusion application, we consider the social cost to be a centralized cost, i.e. the aggregate of the follower costs. Hence, the pricing mechanism is used to close the gap between the centralized and decentralized costs. The pricing mechanisms we propose in both scenarios are designed to integrate seamlessly into existing control schemes without requiring large infrastructural changes.

In Section II, we formally define the pricing mechanism design problem in both the continuous and discrete time framework. In Section III, we state and prove our main theorems. In Section IV, we propose two objective functions for improved performance. In Section V, we apply the theory developed to the problem of network security. In Section VI, we formulate the problem of pricing for controlled diffusion in a general multi-agent network. In Section VII, we make concluding remarks.

Lillian J. Ratliff, Samuel Coogan, Daniel Calderone, and S. Shankar Sastry are with the Department of Electrical Engineering & Computer Sciences at the University of California, Berkeley, California 94720, {ratliffl, scoogan, danjc, sastry}@eecs.berkeley.edu

The work presented is supported by NSF graduate fellowships, NSF CPS:Large:ActionWebs award number 0931843, TRUST (Team for Research in Ubiquitous Secure Technology) which receives support from the NSF (CCF-0424422), ARO MURI SWARMS (W911NF-0510219), and AFOSR MURI CHASE award number FA9550-10-1-0567

II. Problem Formulation

Consider a dynamical game with one leader and p followers. Let $x_i \in \mathbb{R}^{n_i}$, denote the state of the ith follower and let $u_i \in \mathbb{R}^{m_i}$ denote the control input of the ith follower, where n_i is the state dimension and m_i is the input dimension of the ith player. Further, define $n \triangleq \sum_{i=1}^p n_i$, $m \triangleq \sum_{i=1}^p m_i$, and

$$x \triangleq \begin{bmatrix} x_1^T & \dots & x_p^T \end{bmatrix}^T, \qquad u \triangleq \begin{bmatrix} u_1^T & \dots & u_p^T \end{bmatrix}^T. \quad (1)$$

The states are dynamically coupled and evolve under one of the following dynamics:

$$\dot{x}(t) = Ax(t) + \textstyle\sum_{i=1}^p B_i u_i(t) \quad \text{(CT)}$$

or

$$x[k+1] = Ax[k] + \textstyle\sum_{i=1}^p B_i u_i[k] \quad \text{(DT)}$$

for some matrix $A \in \mathbb{R}^{n\times n}$ and $B_i \in \mathbb{R}^{n\times m_i}$ for $i = 1,\dots,p$ and where CT and DT denote continuous and discrete time respectively. When the time dependence is obvious or we wish to discuss both CT and DT states and inputs, we sometimes omit the explicit dependence on the time variable t or k. We make the following assumption:

Assumption 1. *$(A, \begin{bmatrix} B_1 & \cdots & B_p \end{bmatrix})$ is stabilizable.*

Each follower incurs the quadratic cost J_i given by one of the following:

- Continuous-time, infinite horizon cost (CT cost):
$$J_i \triangleq \textstyle\int_0^\infty x(t)^T Q_i x(t) + u(t)^T R_i u(t)\ dt \quad (2)$$
- Discrete time, infinite horizon cost (DT cost):
$$J_i \triangleq \textstyle\sum_{k=0}^\infty x[k]^T Q_i x[k] + u[k]^T R_i u[k] \quad (3)$$
- Discrete-time, finite horizon cost (DTFH cost):
$$\begin{aligned} J_i \triangleq & x[M]^T Q_i[M] x[M] \\ & + \textstyle\sum_{k=0}^{M-1} x[k]^T Q_i[k] x[k] + u[k]^T R_i[k] u[k] \end{aligned} \quad (4)$$

where for CT or DT costs (resp. DTFH costs), for all i:

- $Q_i \in \mathbb{R}^{n\times n}$, $Q_i \succeq 0$ (resp. $Q_i[k] \in \mathbb{R}^{n\times n}$, $Q_i[k] \succeq 0$) is a fixed matrix (resp. are fixed matrices for $k = 0,\dots,M$), and
- $R_i \in \mathbb{R}^{m\times m}$ (resp. $R_i[k] \in \mathbb{R}^{m\times m}$), is a fixed, symmetric matrix (resp. are fixed, symmetric matrices for $k = 0,\dots,M-1$) to be designed by the leader.

The goal is to design a set of matrices $\{R_i\}_{i=1}^p$ or $\{R_i[1],\dots,R_i[M-1]\}_{i=1}^p$ such that the Nash equilibrium of the follower game meets the leader's objective. In the DTFH case, the Nash equilibrium in the space of closed loop, no memory strategies has been shown to be unique [8]. We are primarily interested in applications in which the leader can influence the price of control, but not the preference on states. Hence, we fix Q_i and search for an appropriate R_i for each follower.

We assume that the CT cost is used for games with CT dynamics, and either the DT or the DTFH cost is used for games with DT dynamics. For notation purposes, we partition R_i in the CT and DT cost and write

$$R_i = \begin{bmatrix} R_i^{11} & R_i^{12} & \cdots & R_i^{1p} \\ R_i^{21} & \ddots & & R_i^{2p} \\ \vdots & & \ddots & \vdots \\ R_i^{p1} & \cdots & \cdots & R_i^{pp} \end{bmatrix} \quad (5)$$

with $R_i^{jk} \in \mathbb{R}^{m_j\times m_k}$. Note that, since R_i is symmetric, it must be that $R_i^{jk} = (R_i^{kj})^T$ for all j,k. We similarly partition $R_i[k]$ for each $k = 0,\dots,M-1$ for the DTFH cost.

The leader's only influence on the game is to choose matrices R_i for $i = 1,\dots,p$ or matrices $R_i[k]$ for $i = 1,\dots,p$, $k = 0,\dots,M-1$, and we consider the leader's strategy space Γ_L to be defined as follows:

- If CT or DT costs,
$$\Gamma_L \triangleq \{\{R_i\}_{i=1}^p : R_i \in \mathbb{R}^{m\times m}, R_i^{ii} \succ 0\ i = 1,\dots,p\}. \quad (6)$$
- If DTFH costs,
$$\begin{aligned} \Gamma_L \triangleq \{\{R_i[0],\dots,R_i[M-1]\}_{i=1}^p : R_i[k] \in \mathbb{R}^{m\times m}, \\ R_i^{ii} \succ 0\ i = 1,\dots,p, k = 0,\dots,M-1\}. \end{aligned} \quad (7)$$

The game is played as follows: The leader announces a strategy $\gamma_L \in \Gamma_L$, and each follower responds by choosing a causal feedback control law u_i (the objective of the leader is described subsequently). Each follower's strategy space Γ_i is defined to be the set of all causal, memoryless state feedback controls. We refer to the game played by the followers after the leader announces $\gamma_L \in \Gamma_L$ as the *follower game*. Given γ_L, we assume the followers are rational and collectively play a Nash equilibrium of the follower game. A *Nash equilibrium* is defined as a collection of follower actions $(u_1^*,\dots,u_p^*)$ with $u_i^* \in \Gamma_i$ such that the following holds:

$$J_i(u_i^*, u_{-i}^*) \le J_i(u_i, u_{-i}^*) \quad \forall u_i \in \Gamma_i,\ \forall i \quad (8)$$

where u_{-i}^* denotes the set of actions taken by players other than player i, i.e. $-i \triangleq \{1,\dots,i-1,i+1,\dots,p\}$ and $u_{-i}^* \triangleq \{u_j^*\}_{j\in -i}$. When the followers play a Nash equilibrium, no single player can achieve a lower cost by unilaterally changing his or her strategy.

Now that we have defined the followers' actions in response to a given strategy of the leader, we describe how the leader behaves. We assume that leader has a desired set of feedback controllers:

- If CT costs,
$$u_i^d(t) = -K_i^{\mathrm{d}} x(t) \text{ for } i = 1,\dots,p. \quad (9)$$
- If DT costs,
$$u_i^d[k] = -K_i^{\mathrm{d}} x[k] \text{ for } i = 1,\dots,p. \quad (10)$$
- If DTFH costs,
$$u_i^d[k] = -K_i^{\mathrm{d}}[k] x[k] \text{ for } k = 0,\dots,M-1, i = 1,\dots,p \quad (11)$$

with $K_i^{\mathrm{d}} \in \mathbb{R}^{m_i \times n}$ or $K_i^{\mathrm{d}}[k] \in \mathbb{R}^{m_i \times n}$ for all $k = 1, \dots, M-1$.

In addition, we make the following assumption:

Assumption 2. *The desired feedback gains $\{K_i^d\}_{i=1}^p$ or $\{K_i^d[0], \dots, K_i^d[M-1]\}_{i=1}^p$ stabilize the system.*

Remark 1. *We do not make any assumptions on how the leader obtains the desired feedback gains, but a number of possibilities exist. For example, in the case of CT costs, each agent may be equipped with a nominal $\hat{R}^i$ such that the nominal cost is*

$$\hat{J}_i = \int_0^\infty x(t)^T Q_i x(t) + u(t)^T \hat{R}_i u(t)\ dt, \tag{12}$$

and the leader's cost is then $\sum_{i=1}^p \hat{J}_i$ (analogously for DT and DTFH costs). The leader could then solve the resulting team problem using standard linear quadratic regulator theory (LQR) to obtain $\{K_i^d\}_{i=1}^p$ or $\{K_i^d[0], \dots, K_i^d[M-1]\}_{i=1}^p$, in which case Assumption 2 is guaranteed to be satisfied. Additionally, the leader could choose K_i^d to satisfy some sparsity constraint. For example, the leader could choose K_i^d such that $K_i^d x$ is only a function of x_i.

Remark 2. *If agents do not have access to the full state vector for feedback but instead only have access to $y_i = C_i x_i$, we consider the leader to have desired controllers $u_i = -\tilde{K}_i^d y_i$ or $u_i[k] = -\tilde{K}_i^d[k] y_i[k]$. We can then define $K_i^d = \tilde{K}_i^d C_i$ or $K_i^d[k] = \tilde{K}_i^d[k] C_i$ and proceed as in the full information case.*

Knowing that the followers are rational, the leader's task is to find a strategy $\gamma_L \in \Gamma_L$ such that the followers choose $u_i^* = -K_i^{\mathrm{d}} x$ or $u_i^*[k] = -K_i^{\mathrm{d}}[k]x[k]$, i.e. $\{u_i^*\}_{i=1}^p$ is a Nash equilibrium of the follower game.

We first review standard results from LQR theory [9], [10]:

Proposition 1. *Consider the dynamical system $\dot{x} = Ax + Bu$ with (A, B) stabilizable. If for $R \succ 0$, $\begin{bmatrix} Q & N \\ N^T & R \end{bmatrix} \succ 0$ there exists a matrix $P \succeq 0$ such that*

$$A^T P + PA + Q - (PB + N)R^{-1}(B^T P + N^T) = 0, \tag{13}$$

then the feedback law $u = -Kx$ with

$$K = R^{-1}(B^T P + N^T) \tag{14}$$

minimizes

$$J = \int_0^\infty x^T Q x + u^T R u + 2x^T N u\ dt \tag{15}$$

and $A - BR^{-1}(B^T P + N^T)$ is stable. The minimizing cost is $J^* = x(0)^T P x(0)$.

Note that (13) is equivalent to

$$A^T P + PA + Q - K^T R K = 0. \tag{16}$$

Proposition 2. *Consider the dynamical system $x[k+1] = Ax[k] + Bu[k]$ with (A, B) stabilizable. If for $R \succ 0$, $\begin{bmatrix} Q & N \\ N^T & R \end{bmatrix} \succ 0$ there exists a matrix $P \succeq 0$ such that*

$$P = Q + A^T P A - K^T (B^T P B + R) K \tag{17}$$

with

$$K = (B^T P B + R)^{-1}(B^T P A + N^T), \tag{18}$$

then the feedback law $u = -Kx$ minimizes

$$J = \textstyle\sum_{k=0}^\infty x[k]^T Q x[k] + u[k]^T R u[k] + 2x[k]^T N u[k] \tag{19}$$

and $A - BK$ is stable. The minimizing cost is $J^* = x[0]^T P[0] x[0]$.

Define the discrete time index set $\mathcal{K} := \{0, \dots, M-1\}$.

Proposition 3. *Consider the dynamical system $x[k+1] = A[k]x[k] + B[k]u[k]$. Assume $R[k] \succ 0$ for all $k \in \mathcal{K}$, $\begin{bmatrix} Q[k] & N[k] \\ N[k]^T & R[k] \end{bmatrix} \succ 0$ for all $k \in \mathcal{K}$, and $Q[M] \succ 0$. If there exists matrices $P[k] \succeq 0$ such that $P[M] = Q[M]$ and*

$$\begin{aligned} P[k] &= Q[k] + A[k]^T P[k+1] A[k] \\ &\quad - K[k]^T (B[k]^T P[k+1] B[k] + R[k]) K[k], \quad \forall k \in \mathcal{K} \end{aligned} \tag{20}$$

with

$$\begin{aligned} K[k] &= (B[k]^T P[k+1] B[k] + R[k])^{-1} \\ &\quad \cdot (B[k]^T P[k+1] A[k] + N[k]^T), \quad \forall k \in \mathcal{K} \end{aligned} \tag{21}$$

then the feedback law $u[k] = -K[k]x[k]$, $\forall k \in \mathcal{K}$ minimizes

$$\begin{aligned} J = &\ x[M]^T Q[M] x[M] + \textstyle\sum_{k=0}^{M-1} \big(x[k]^T Q[k] x[k] \\ &+ u[k]^T R[k] u[k] + 2x[k]^T N[k] u[k] \big). \end{aligned} \tag{22}$$

The minimizing cost is $J^* = x[0]^T P[0] x[0]$.

We can apply Propositions 1–3 to develop a methodology for obtaining $\gamma_L \in \Gamma_L$ that achieves the leader's goal.

III. Main Results

In the analysis that follows, we investigate player i's optimal control assuming that all other players use strategies $u_j = -K_j^{\mathrm{d}} x$ or $u_j[k] = -K_j^{\mathrm{d}}[k] x[k]$. We justify this assumption by noting that the class of feedback strategies are strongly time consistent, [11]. One may compare this with open-loop strategies which are weakly time consistent since player strategies are dependent on the initial state only. Hence, any deviation by a player from his initial strategy may result in the open-loop solution being suboptimal. In addition, since we restrict to the class of feedback strategies, if one player uses feedback then it is optimal for the others to do so as well based on the strong time consistency property and the rationality of the players.

We investigate separately the cases of continuous time dynamics with infinite horizon costs, discrete time dynamics with infinite horizon costs, and discrete-time dynamics with finite horizon costs.

A. Continuous-Time, Infinite Horizon

Assume CT dynamics with CT costs. We introduce the following notation. Let R_i^{-i-i} denote R_i with the ith block-row and the ith block-column removed. Similarly, let R_i^{i-i} and R_i^{-ii} denote the ith block-row and the ith block-column of R_i each with the R_i^{ii} block removed. Note that $R_{i-i}^i =$

$(R_i^{-ii})^T$ By abuse of notation (i.e., after an appropriate row/column permutation), we have

$$R_i \triangleq \begin{bmatrix} R_i^{ii} & R_i^{i-i} \\ R_i^{-ii} & R_i^{-i-i} \end{bmatrix}. \tag{23}$$

We also introduce the following notation for the leader's desired control for all the followers except for follower i:

$$K_{-i}^{\mathrm{d}} = \begin{bmatrix} K_1^{\mathrm{d}T} & \cdots & K_{i-1}^{\mathrm{d}}{}^T & K_{i+1}^{\mathrm{d}}{}^T & \cdots & K_p^{\mathrm{d}T} \end{bmatrix}^T. \tag{24}$$

Furthermore, define

$$\tilde{A}_i \triangleq A - \textstyle\sum_{j\neq i} B_j K_j^{\mathrm{d}} = A - B_{-i}K_{-i}^{\mathrm{d}} \tag{25}$$

$$\begin{aligned}\tilde{Q}_i &\triangleq Q_i + \textstyle\sum_{j\neq i,\ell\neq i}(K_j^{\mathrm{d}})^T R_i^{j\ell} K_\ell^{\mathrm{d}} \\ &= Q_i + (K_{-i}^{\mathrm{d}})^T R_i^{-i-i} K_{-i}^{\mathrm{d}}\end{aligned} \tag{26}$$

$$\tilde{N}_i \triangleq -\textstyle\sum_{j\neq i}(K_j^{\mathrm{d}})^T R_i^{ji} = -(K_{-i}^{\mathrm{d}})^T R_i^{-ii} \tag{27}$$

for all i. Note that $(\tilde{A}_i, B_i)$ is stabilizable by Assumption 2.

Under the assumption that all players $j \neq i$ choose $u_j = -K_j^{\mathrm{d}}x$, player i will experience the dynamical system

$$\dot{x} = \tilde{A}_i x + B_i u_i \tag{28}$$

with cost

$$J_i = \textstyle\int_0^\infty x^T\tilde{Q}_i x + u_i^T R_i^{ii} u_i + 2x^T \tilde{N}_i u_i \; dt. \tag{29}$$

We can now apply Proposition 1 and obtain the following:

Theorem 1 (CT costs)**.** *If $\gamma^* = \{R_i\}_{i=1}^p$ and $\{P_i\}_{i=1}^p$ exists such that the convex feasibility problem below is feasible for all $i \in \{1,\ldots,p\}$, then $\{u_i^* = u_i^d = -K_i^d x\}_{i=1}^p$ is a Nash equilibrium of the follower game, thereby achieving the leader's goal:*

$$P_i \succeq 0 \tag{30}$$

$$R_i^{ii} \succ 0 \tag{31}$$

$$\begin{bmatrix} \tilde{Q}_i & \tilde{N}_i \\ (\tilde{N}_i)^T & R_i^{ii} \end{bmatrix} \succeq 0 \tag{32}$$

$$(\tilde{A}_i)^T P_i + P_i(\tilde{A}_i) + \tilde{Q}_i - (K_i^d)^T R_i^{ii} K_i^d = 0 \tag{33}$$

$$(B_i^T P_i + (\tilde{N}_i)^T) = R_i^{ii} K_i^d \tag{34}$$

Proof: By (8), $\{u_i^d = -K_i^d x\}_{i=1}^p$ is a Nash equilibrium for the follower game if for each i, $u_i = -K_i^d x$ is optimal for cost (29) subject to dynamics (28). If there exists R_i, P_i that satisfy (30)–(34) for all i, then by Proposition (1), $u_i^d = -K_i^d x$ is optimal for (29) for all i. ■

B. Discrete-Time, Infinite Horizon

Assume DT dynamics and DT costs, and consider definitions (23)–(27). Under the assumption that all players $j \neq i$ choose $u_j[k] = -K_j^{\mathrm{d}}x[k]$, player i will experience the dynamical system

$$x[k+1] = \tilde{A}_i x[k] + B_i u_i[k] \tag{35}$$

with cost

$$J_i = \textstyle\sum_{k=0}^\infty x[k]^T\tilde{Q}_i x[k] + u_i[k]^T R_i^{ii} u_i[k] + 2x[k]^T\tilde{N}_i u_i[k]. \tag{36}$$

Applying Proposition 2, we have the following:

Theorem 2. *If $\gamma^* = \{R_i\}_{i=1}^p$ and $\{P_i\}_{i=1}^p$ exists such that the convex feasibility problem below is feasible for all $i \in \{1,\ldots,p\}$, then $\{u_i^*[k] = u_i^d[k] = -K_i^d x[k]\}_{i=1}^p$ is a Nash equilibrium of the follower game, thereby achieving the leader's goal:*

$$P_i \succeq 0 \tag{37}$$

$$R_i^{ii} \succ 0 \tag{38}$$

$$\begin{bmatrix} \tilde{Q}_i & \tilde{N}_i \\ (\tilde{N}_i)^T & R_i^{ii} \end{bmatrix} \succeq 0 \tag{39}$$

$$\tilde{Q}_i + (\tilde{A}_i)^T P_i \tilde{A}_i - P_i \tag{40}$$

$$-(K_i^d)^T(B_i^T P_i B_i + R_i^{ii})K_i^d = 0 \tag{41}$$

$$(B_i^T P_i B_i + R_i^{ii})K_i^d = (B_i^T P_i \tilde{A}_i + (\tilde{N}_i)^T) \tag{42}$$

The proof is analogous to the proof of Theorem 1 and is omitted.

C. Discrete-Time, Finite Horizon

Assume DT dynamics and DTFH costs. Analogous to the previous cases, define:

$$R_i[k] \triangleq \begin{bmatrix} R_i^{ii}[k] & R_i^{i-i}[k] \\ R_i^{-ii}[k] & R_i^{-i-i}[k] \end{bmatrix} \tag{43}$$

$$K_{-i}^{\mathrm{d}}[k] \triangleq \begin{bmatrix} K_1^{\mathrm{d}}[k]^T \cdots K_{i-1}^{\mathrm{d}}[k]^T & K_{i+1}^{\mathrm{d}}[k]^T \cdots K_p^{\mathrm{d}}[k]^T \end{bmatrix}^T \tag{44}$$

$$\tilde{A}_i[k] \triangleq A - B_{-i}K_{-i}^{\mathrm{d}}[k] \tag{45}$$

$$\tilde{Q}_i[k] \triangleq Q_i + (K_{-i}^{\mathrm{d}}[k])^T R_i^{-i-i}[k] K_{-i}^{\mathrm{d}}[k] \tag{46}$$

$$\tilde{N}_i[k] \triangleq -(K_{-i}^{\mathrm{d}}[k])^T R_i^{-ii}[k]. \tag{47}$$

Under the assumption that all players $j \neq i$ choose $u_j[k] = -K_j^{\mathrm{d}}[k]x[k]$, player i will experience the dynamical system

$$x[k+1] = \tilde{A}_i x[k] + B_i u_i[k] \tag{48}$$

with costs

$$\begin{aligned} J_i = x[M]^T\tilde{Q}_i[M]x[M] + \textstyle\sum_{k=0}^M \Big(x[k]^T\tilde{Q}_i[k]x[k] \\ + u_i[k]^T R_i^{ii}[k]u_i[k] + 2x[k]^T\tilde{N}_i[k]u_i[k] \Big). \end{aligned} \tag{49}$$

Applying Proposition 3, we have the following:

Theorem 3. *If $\gamma^* = \{R_i[0],\ldots,R_i[M-1]\}_{i=1}^p$ and $\{P_i[0],\ldots,P_i[M]\}_{i=1}^p$ exists such that the convex feasibility problem below is feasible for all $i \in \{1,\ldots,p\}$, then $\{u_i^*[k] = u_i^d[k] = -K_i^d[k]x[k]\}_{i=1}^p$ is the Nash equilibrium of the follower game, thereby achieving the leader's goal:*

$$P_i[M] = \tilde{Q}_i[M] \tag{50}$$

$$R_i^{ii}[k] \succ 0 \tag{51}$$

$$\begin{bmatrix} \tilde{Q}_i[k] & \tilde{N}_i[k] \\ (\tilde{N}_i[k])^T & R_i^{ii}[k] \end{bmatrix} \succeq 0 \tag{52}$$

$$\begin{aligned} P_i[k] =& \tilde{Q}_i[k] + \tilde{A}_i[k]^T P_i[k+1]\tilde{A}_i[k] \\ &- (K_i^d[k])^T(B_i^T P_i[k+1]B_i + R_i^{ii}[k])K_i^d[k] \end{aligned} \tag{53}$$

$$(B_i^T P_i[k] B_i + R_i^{ii}[k]) K_i^d[k] = (B_i^T P_i[k] \tilde{A}_i[k] + (\tilde{N}_i[k])^T) \quad (54)$$

where (51)-(54) hold for all $k \in \mathcal{K}$.

The proof is analogous to the proof of Theorem 1 and is omitted. Observe that we have considered the problem of finding prices for the ith follower as an inverse LQR problem with the additional constraint that Q^i is fixed.

IV. Objective Functions for Improved Performance

Often, the feasibility problems from Theorem 1–3 have many solutions, thereby providing the opportunity to add objectives for finding a desirable $\gamma_L^* \in \Gamma_L$. We discuss two possibilities below.

A. Revenue Neutral Pricing

By considering the nominal costs with associated prices $\hat{\gamma}_L$, the leader calculates *team optimal* or *socially optimal* control inputs and, using these as the desired controllers, the leader finds $\gamma_L^* \in \Gamma_L$ via Theorems 1–3 such that the team optimal controller is a Nash equilibrium for the follower game. Thus, the difference between costs incurred under γ_L^* versus $\hat{\gamma}_L$ can be interpreted as *incentives* or *penalties* established by the leader to induce the team optimal solution.

Let $\hat{R}_i$, $\hat{J}_i$, etc. be associated with the game with nominal leader strategy $\hat{\gamma}_L$. There is no *a priori* association between the nominal incurred costs $\hat{J}_i$ and the costs after incentive pricing, J_i^*. However, it may be desirable to coerce a relationship either using an objective function or additional problem constraints. For instance, if $\sum_{i=1}^p J_i^* < \sum_{i=1}^p \hat{J}_i$, the costs "collected" by the leader are less than the nominal costs that are incurred. Conversely, if $\sum_{i=1}^p J_i^* \gg \sum_{i=1}^p \hat{J}_i$, the followers may find the game unduly unfair.

Now consider the feasibility problems of Theorems 1–3. Define

$$\mathcal{S}_1 := \{(\{R_i\}_{i=1}^p, \{P_i\}_{i=1}^p) : \text{(30)-(34) hold}\}, \quad (55)$$

$$\mathcal{S}_2 := \{(\{R_i\}_{i=1}^p, \{P_i\}_{i=1}^p) : \text{(37)-(42) hold}\}, \text{ and} \quad (56)$$

$$\begin{aligned}\mathcal{S}_3 := \{(&\{R_i[0], \ldots, R_i[M-1]\}_{i=1}^p,\\ &\{P_i[0], \ldots, P_i[M]\}_{i=1}^p) : \text{(51)-(54) hold}\}. \end{aligned} \quad (57)$$

Suppose we wish $\sum_{i=1}^p J_i^* = \sum_{i=1}^p \hat{J}_i$. One solution is to enforce $\sum_{i=1}^p R_i = \sum_{i=1}^p \hat{R}_i$, or $\sum_{i=1}^p R_i[k] = \sum_{i=1}^p \hat{R}_i[k]$ for all k. This ensures that the total follower costs of the modified game at each time step is equal to the total nominal follower costs at each time step. However, this is often an unattainable goal. We alternatively consider $\sum_{i=1}^p J_i$ versus $\sum_{i=1}^p \hat{J}_i$, which compares the total costs throughout the modified game versus the total nominal costs. Ideally, we want $\sum_{i=1}^p J_i = \sum_{i=1}^p \hat{J}_i$, however we can instead add the relaxed *revenue neutral* objective $\min \left|\sum_{i=1}^p J_i - \sum_{i=1}^p \hat{J}_i\right|$. We can now formulate the following optimization problems given the *revenue neutral* objective:

$$\begin{aligned}&\min |\textstyle\sum_i \mathbb{E}(x(0)^T P_i^0 x(0)) - C| \\ &\text{subj. to } (\{R_i\}_{i=1}^p, \{P_i\}_{i=1}^p) \in S_o\end{aligned} \quad \text{(P-1)}$$

where $o = 1, 2, 3$ if the problem is CT, DT, or DTFH respectively, and where $P_i^0 = P_i$ if the problem is CT or DT and $P_i^0 = P_i[0]$ if the problem is DTFH. If $x(0)$ is known, then $\mathbb{E}(x(0)^T P_i^0 x(0)) = x(0)^T P_i^0 x(0)$ and $C = \sum_{i=1}^p \hat{J}_i$. If $x(0)$ is not known *a priori* we can assume a probability distribution on $x(0)$ and use expectation. In this case, $C \triangleq \mathbb{E}(\sum_{i=1}^p \hat{J}_i)$. A standard assumption is that $x(0)$ is uniformly distributed on the unit sphere, in which case $\mathbb{E}(x(0)^T P_i^0 x(0)) = \mathbf{Tr}(P_i^0)$ where the expectation is taken over $x(0)$.

We can also easily include constraints such as $\sum_{i=1}^p x(0)^T P_i x(0) - C \geq 0$ or $\sum_{i=1}^p \mathbf{Tr}(P_i) - C \geq 0$, and similarly for DTFH costs to ensure that the total modified costs are at least the total nominal costs. In simulations and examples, it is often the case that prices can be found such that $\sum_{i=1}^p J_i = \sum_{i=1}^p \hat{J}_i$.

B. Minimizing Impact of Non-Local Controls on Cost

In general, solutions to the feasibility problems from Theorems 1–3 produce pricing mechanisms for specific agents that depend on other agents control actions as well as their own control actions. In many situations, this kind of pricing mechanism could be construed as unfair by competitive agents who do not want to be penalized for the actions of their neighbors. In any market economy, the price of a scarce resource for a particular agent depends on how much of that resource is being used by other agents; however, this may do little to dispel the perception of unfairness. In order to mitigate this problem, the leader may design pricing mechanisms for each agent that have minimal dependence on other agents' control actions while still inducing the agents to use the desired feedback gains. This amounts to minimizing the norm of the R_i^{i-i}, R_i^{-ii}, and R_i^{-i-i} blocks in Equation (23). We can now formulate the following optimization problems:

$$\min \textstyle\sum_{j\neq i, \ell\neq i} w_{j\ell}^i \|R_i^{j\ell}\|_2 + \sum_{j\neq i} w_{ij}^i \left(\|R_i^{ij}\|_2 + \|R_i^{ji}\|_2\right) \quad \text{(P-2)}$$

$$\text{subj. to } (\{R_i\}_{i=1}^p, \{P_i\}_{i=1}^p) \in \mathcal{S}_o,$$

$$\begin{aligned}\min \textstyle\sum_{k=1}^{M-1} \Big(&\textstyle\sum_{j\neq i, \ell\neq i} w_{j\ell}^i \|R_i^{j\ell}[k]\|_2 \\ &+ \textstyle\sum_{j\neq i} w_{ij}^i \left(\|R_i^{ij}[k]\|_2 + \|R_i^{ji}[k]\|_2\right)\Big)\end{aligned} \quad \text{(P-3)}$$

$$\begin{aligned}\text{subj. to } (&\{R_i[0], \ldots, R_i[M-1]\}_{i=1}^p,\\ &\{P_i[0], \ldots, P_i[M]\}_{i=1}^p) \in \mathcal{S}_3.\end{aligned}$$

where in (P-2) $o = 1, 2$ if the problem is CT or DT respectively. We call these objectives *local pricing* objectives.

The weighting terms $w_{j\ell}^i$ allow the leader to place more emphasis on minimizing specific subblocks of R_i. For example, the leader might care about minimizing the R_i^{-i-i} block more since it is the part of the pricing mechanism that depends solely on other agents' control actions whereas the R_i^{i-i} and R_i^{-ii} blocks depend both on agent i's actions and other agents' actions.

In addition to considering the *revenue neutral* and *local pricing* objectives individually, we may consider a weighted sum of the two objectives with weights w_{rn} and w_{lp} for the *revenue neutral* and *local pricing* objectives respectively.

V. Example 1: Pricing for Network Security

In this section we consider a numerical example of the pricing mechanism design problem for coordinating multiple interconnected networks to invest in security. Self-spreading attacks on computer networks are expensive owing to the damage they cause and the security investment required to defend against them [12], [13]. The leader's goal is to design pricing mechanisms that coordinate the networks so that the overall multi-network is stabilized.

A. Epidemic Model for Multi-Network

A common model for self-spreading attacks is the classical epidemic model adapted to a multi-network framework. In the multi-network epidemic model the state x_i denotes the fraction of infected hosts in network i for $i = 1, \dots, p$. Let the control action u_i be the malware removal rate for network i. The parameters α and β are the cross-network and inter-network pairwise infection rates respectively. Let N_i denote the number of hosts in network i for $i = 1, \dots, p$. We may analyze a modified version of the nonlinear epidemic model given in [12] since the modified model is more difficult to stabilize. The modified model is given by

$$\dot{x}_i = \begin{cases} [Ax + Bu]_i, & (x_i > 0) \vee ([Ax + Bu]_i \geq 0 \wedge x_i = 0) \\ 0, & \text{otherwise} \end{cases} \tag{58}$$

where $A \in \mathbb{R}^{p \times p}$ with entries

$$A_{i,j} = \left\{ \begin{array}{ll} \beta N_i, & \text{if } \ i = j \\ \alpha N_i, & \text{otherwise.} \end{array} \right. \tag{59}$$

and $B = -I_{p \times p}$.

B. Network Utilities

Each network independently tries to choose u_i so that the ith network is stabilized. We propose the use of pricing mechanisms to coordinate the networks by inducing them to choose a desired control action which stabilizes the entire multi-network. In Bloem, et. al., Theorem 1 of [13], it is proven that a system with dynamics (58) is stable with feedback $u = -K^{\text{d}}x$ if the closed-loop matrix has negative diagonal entries and non-positive off-diagonal entries. As such, we define the desired solution $\{K_i^{\text{d}}\}_{i=1}^p$ to be a set of feedback gains that stabilize the closed loop system in the sense of this theorem.

The individual networks each have quadratic nominal utilities given by

$$\hat{J}_i = \int_0^\infty x_i^T Q_i x_i + u_i^T \hat{R}_i u_i \ dt \tag{60}$$

where Q_i and $\hat{R}_i$ is the cost of an infected network host and the cost of the implemented patching response rate for network i respectively. The leader chooses a set of pricing matrices $\{R_i\}_{i=1}^p$ such that network i is induced to choose the stabilizing feedback gain K_i^{d} for $i = 1, \dots, p$.

TABLE I
Costs under an epidemic model for network security

	Individual Nominal Costs [a]				CTS[b]	
	N_1	N_2	N_3	N_4		
NOM	1.260	1.288	1.226	1.253	5.028	
TO	1.247	1.279	1.209	1.239	4.974	
	Individual Costs w/ Pricing[c]				CTS	Leader Profit[d]
	N_1	N_2	N_3	N_4		
NO	3.709	3.707	3.709	3.710	4.974	9.860
RN	1.245	1.247	1.238	1.244	4.974	0
LP	1.134	1.133	1.114	1.132	4.974	-0.461
RN, LP	1.193	1.454	1.149	1.178	4.974	0

[a] Nominal costs incurred at the original Nash equilibrium with nominal costs (NOM) and when players use the team optimal control (TO).

[b] Cost To Society, the cost incurred under the nominal costs.

[c] Costs incurred under four different pricing schemes designed with no objectives (NO), with only the revenue neutral objective (RN), with only the local pricing objective (LP), and with both the revenue neutral and local pricing objectives, see section IV. Followers use $\{K_i^{\text{d}}\}_{i=1}^p$.

[d] The difference between the CTS and the total cost with pricing. If the pricing scheme is revenue neutral, the leader profit is 0.

C. Simulation Results

For the numerical simulations we take $\alpha = \frac{2\beta}{3}$ since it is assumed that $\alpha < \beta$ implies attacks will spread at a slower rate if security is in place [13]. We compute the Nash equilibrium of the follower game under the nominal costs using the method of Lyapunov iterations, see [14]. We refer to this cost as the nominal cost. We determine a set of desired feedback gains for the leader by computing the solution to the infinite time LQR optimization problem for the whole multi-network and we use the resulting feedback gains if they satisfy Theorem 1 of [13]. Otherwise, we modify them to satisfy the theorem. We use code written in MATLAB that employs YALMIP to solve the optimization problem[1], [15].

1) Comparing Objectives: For the first set of simulations, we fix the number of networks to be $p = 4$ and let $N_1 = 3500$, $N_2 = 3900$, $N_3 = 3000$ and $N_4 = 3400$. We then design pricing schemes, namely $\{R_i\}_{i=1}^p$, without minimizing any objective, using the revenue neutral objective in (P-1) for CT, using the local-pricing objective in (P-2), and finally using a weighted sum of these two objectives. For simplicity, we let $w_{j\ell}^i = w_{ij}^i = 1$ for all i, j, ℓ and we let $w_{\text{rn}} = w_{\text{lp}} = 1$. Table I records the nominal cost to each network, the team or centralized cost (where networks use $\{K_i^{\text{d}}\}_{i=1}^p$), and the costs under the four pricing schemes.

All four pricing schemes induce each network to use the desired feedback gains but the values of the matrices $\{R_i\}_{i=1}^p$ and the costs incurred by each network under the pricing schemes are different. One should note that if the pricing scheme is designed without an objective, the solver simply returns the first feasible set of pricing matrices to which it converges. In this example, the prices chosen without objectives are significantly higher than the nominal cost incurred by each network at the team optimal solution. The revenue neutral objective minimizes the difference between the team optimal cost and the aggragate of the followers'

[1] Code available at
`www.eecs.berkeley.edu/~scoogan/allerton12.html`

costs under the pricing scheme. In this example, the costs are rendered equal.

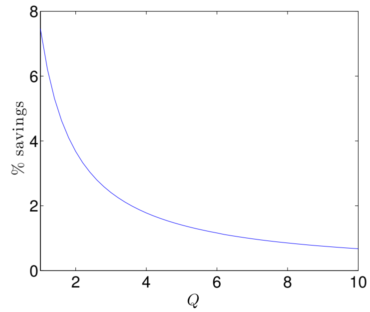

Fig. 1. $p = 4$, percent savings using pricing scheme plotted against $Q : \hat{R}$.

Applying the local pricing objective significantly reduces the magnitude of the non-local terms of the R_1 matrix. However, the magnitude of the non-local terms depends greatly on the weights used in the local pricing objective.

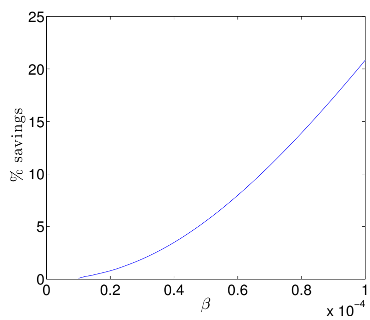

Fig. 2. $p = 4$, percent savings using pricing scheme plotted against β

2) Varying Parameters: For the second simulation, using the same multi-network as in the previous example, we first vary Q_i in the ratio $Q_i : \hat{R}_i$ from 1 to 10 with fixed $\hat{R}_i = 1$ for each i and we fix the inter-network infection rate $\beta = 5.8 \times 10^{-5}$. Then, we vary the inter-network infection rate β from 1×10^{-5} to 1×10^{-4} and we fix $Q_i : \hat{R}_i = 1 : 1$ for all i. In both cases, we minimize a weighted sum of the objectives in (P-1) for CT and (P-2) with $w_{\text{rn}} = 0.5$ and $w_{\text{lp}} = 1$ and we compute the percentage savings gained from using the pricing mechanism as the nominal cost. Figure 1 shows the percentage savings as a function of the cost ratio $Q : \hat{R}$. The percent savings decreases with increasing $Q : \hat{R}$ ratio. This is reasonable since Q represents the cost of the infection to the network. Hence, as the $Q : \hat{R}$ ratio increases it cost more to the network for infections relative to the cost of patching. Figure 2 shows the percentage savings as a function of β. We see that the percent savings increases with β which is reasonable since β represents the strength of the attack.

3) Varying Network Sizes: In the last set of simulations, we consider a multi-network with $p = 6$ networks of two different sizes. We again minimize a weighted sum of the objectives in (P-1) for CT and (P-2) with $w_{\text{rn}} = 1$ and $w_{\text{lp}} = 1$ with respect to $\{R_i\}_{i=1}^p$ and subject to $(\{R_i\}_{i=1}^p, \{P_i\}_{i=1}^p) \in \mathcal{S}_1$. Let the first multi-network have network sizes $N_1 = 5000$, $N_2 = 3000$, $N_3 = 2000$, $N_4 = 900$, $N_5 = 90$, and $N_6 = 1000$. Similarly, for the second multi-netowk let $N_1 = 5000$, $N_2 = 4000$, $N_3 = 4500$, $N_4 = 5500$, $N_5 = 6000$, and $N_6 = 4750$. Table II contains the nominal and the pricing induced cost for each agent as well as the total cost. The first multi-network we consider has one small network with 90 nodes and the other networks are roughly the same size and an order of magnitude larger than the smallest network. In this example, the pricing induced strategies save 9.64% as compared to the nominal cost. The second multi-network contains networks that are all the same size in magnitude and the savings in this example are 37.61%. We conjecture that the difference in percentage savings is exemplary of networks of the same size versus ones with greatly varying sizes.

TABLE II
COST FOR $p = 6$ MULTI-NETWORK EXAMPLES 1 AND 2

Network	**Multi-Network 1** Network Size	Cost w/ Pricing	Nom. Cost	**Multi-Network 2** Network Size	Cost w/ Pricing	Nom. Cost
1	5000	1.355	1.695	5000	1.439	1.969
2	3000	1.293	1.427	4000	1.343	1.815
3	2000	1.190	1.302	4500	1.395	1.891
4	900	1.129	1.173	5500	1.474	2.050
5	90	1.085	1.083	6000	1.499	2.133
6	1000	1.121	1.185	4750	1.418	1.930
Total		7.173	7.865		8.567	11.789
Savings		**9.64%**			**37.61%**	

A further consideration would be to assess the cost of implementing the pricing scheme to the leader. It is interesting to note that in this formulation the leader chooses the desired feedback gains $\{K_i^{\text{d}}\}_{i=1}^p$ simply with the goal of stabilizing the network. Another formulation would be to allow the leader to choose the optimal feedback gains $\{K_i^{\text{d}}\}_{i=1}^p$ that stabilize the network where optimality is defined with respect to some utility representing the cost of implementing the pricing mechanism. We leave this for future work.

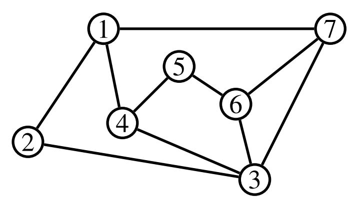

Fig. 3. Network of 7 interconnected players.

VI. EXAMPLE 2: PRICING FOR CONTROLLED DIFFUSION

Consider the network shown in Fig 3 where each agent $i = 1, \dots, 7$ has state $\xi_i[k] \in \mathbb{R}$ at time k. A common control problem is to design a distributed control strategy $f_i[k]$ such that the states converge to the same value under the dynamics $\xi_i[k+1] = f_i[k]$, i.e. $|\xi_i[k] - \xi_j[k]| \to 0$ for all i, j. A solution is the following discretized version of the so-called *consensus* strategy [16]:

$$\xi_i[k+1] = \xi_i[k] - \epsilon \textstyle\sum_{j \in \mathcal{N}_j} (\xi_i[k] - \xi_j[k]) + u_i \qquad (61)$$

where $\mathcal{N}_j = \{j : \text{an edge connects } i \text{ to } j\}$ is the neighborhood of agent i, $\epsilon > 0$ is a chosen stepsize, and we have included an exogenous input u_i. We call problems with an

TABLE III

RESULTS FOR EXAMPLE 2.

Teams	Nominal Cost	Cost with Pricing Mechanism[a]
$\{1\},\{2\},\ldots,\{6\},\{7\}$	11.0785	11.0454
$\{1,2\},\{3,4\},\{5,6,7\}$	11.0638	11.0454
$\{1,2,3,4\},\{5,6,7\}$	11.0561	11.0454
$\{1,\ldots,7\}$(i.e., Team Optimal)	11.0454	N/A

[a]Since the cost with pricing is equal to the team optimal cost, we see that revenue neutral pricing is achieved in all three cases.

exogenous input *controlled diffusion* problems and they arise often in multirobot rendezvous and flocking [17], and sensor networks [18]. We define the matrix L entrywise as

$$[L]_{ij} = \begin{cases} -1 & \text{if } i \neq j \text{ and } i \text{ is connected to } j \\ d_i & \text{if } i = j \end{cases} \tag{62}$$

where d_i is the degree of node i in the network. We let $\xi = \begin{bmatrix}\xi_1 & \ldots & \xi_7\end{bmatrix}^T$, $u = \begin{bmatrix}u_1 & \ldots & u_7\end{bmatrix}^T$, and write the controlled diffusion dynamics as

$$\xi[k+1] = (I_7 - \epsilon L)\xi[k] + u \tag{63}$$

Where I_7 is the 7×7 identity matrix. We consider three noncooperative dynamic games: 1) each agent is a noncooperative player, 2) the agents are divided into 3 noncooperative teams(players), 3) the agents are divided into 2 noncooperative teams(players), see Table III for the team divisions. In addition, we consider the team optimal solution, which is equivalent to one team with all the players. In each case, a player's control input is the set of exogenous inputs to that players' agents. The cost to each player i is assumed to be $\sum_{k=0}^{\infty}(\sum_{j\in T_i}\xi_i^2 + u_i^2)$ where T_i is the indices of agents on team i. A full state optimal controller is found by solving the LQR problem $\min \sum_{k=0}^{\infty} \xi^T\xi + u^T u$. For a given team, we assume that the states of all the agents on that team and the states of their neighbors are available for measurement, and suboptimal desired controllers for each player are then obtained from the full state optimal controller by extracting the appropriate rows for each player and zeroing out entries corresponding to state measurements unavailable to that player. We then find new prices for the game using Theorem 2 and a revenue neutral objective. Results are summarized in Table III, where we choose $\epsilon = 0.1$. We see that exactly revenue neutral prices are possible in all three cases considered.

VII. CONCLUSION

This paper investigated pricing for linear-quadratic dynamic games. We posed the game as a dynamic game with many followers and one leader. We used results from LQR theory to derive our main result showing that the problem of obtaining costs to ensure a particular desired Nash equilibrium for the followers is a convex feasibility problem. We presented extensions to this problem which maintain the convexity of the resulting program, namely making costs revenue neutral and minimizing dependence of costs on nonlocal control actions.

Future work involves examining the robustness of the pricing scheme as well as the Nash equilibrium strategies to affine perturbations and characterization of the feasible set of pricing matrices. In addition, a related problem to pricing for controlled diffusion is incentivizing efficient HVAC control and occupant behavior in buildings. In future research, we will consider a HVAC system model proposed in [19] and design pricing schemes with the goal of improving overall efficiency.

VIII. ACKNOWLEDGEMENTS

The authors thank Murat Arcak, Anil Aswani, Andrew Lim, Claire Tomlin, and Ram Vasudevan for valuable discussions.

REFERENCES

[1] T. Chang and Y. Ho, "Incentive problems: A class of stochastic stackelberg closed-loop dynamic games," *Systems & Control Letters*, vol. 1, no. 1, pp. 16–21, 1981.

[2] Y. Ho, P. Luh, and G. Olsder, "A control-theoretic view on incentives," *Automatica*, vol. 18, no. 2, pp. 167–179, 1982.

[3] J. Laffont and D. Martimort, *The Theory of Incentives I: The Principal-Agent Model*. Princeton University Press, 2002.

[4] M. Pachter, "Linear-quadratic reversed Stackelberg differential games with incentives," *IEEE Transactions on Automatic Control*, vol. 29, no. 7, pp. 644–647, 1984.

[5] J. Marden and A. Wierman, "Distributed welfare games with applications to sensor coverage," in *Proceedings of the 47th IEEE Conference on Decision and Control*, Dec. 2008, pp. 1708–1713.

[6] ——, "Overcoming limitations of game-theoretic distributed control," in *Proceedings of the 48th IEEE Conference on Decision and Control, 2009*, Dec. 2009, pp. 6466–6471.

[7] N. Li and J. Marden, "Designing games to handle coupled constraints," in *Proceedings of the 49th IEEE Conference on Decision and Control*. IEEE, 2010, pp. 250–255.

[8] T. Basar, "On the uniqueness of the nash solution in linear-quadratic differential games," *International Journal of Game Theory*, vol. 5, pp. 65–90, 1976, 10.1007/BF01753310. [Online]. Available: http://dx.doi.org/10.1007/BF01753310

[9] J. P. Hespanha, *Linear Systems Theory*. Princeton University Press, 2009.

[10] D. Bertsekas, *Dynamic Programming and Optimal Control*, 2nd ed. Athena Scientific, 2005, vol. 1.

[11] T. Basar and G. J. Olsder, *Dynamic Noncooperative Game Theory*, 2nd ed., ser. Classics in Applied Mathematics. SIAM, 1999.

[12] T. Alpcan and T. Başar, *Network Security: A Decision and Game-Theoretic Approach*. Cambridge Univ Press, 2010.

[13] M. Bloem, T. Alpcan, and T. Basar, "Optimal and robust epidemic response for multiple networks," *Control Engineering Practice*, vol. 17, no. 5, pp. 525–533, 2009.

[14] Z. Gajić and M. Qureshi, *Lyapunov matrix equation in system stability and control*. Academic Pr, 1995, vol. 195.

[15] J. Löfberg, "Yalmip : A toolbox for modeling and optimization in MATLAB," in *Proceedings of the CACSD Conference*, Taipei, Taiwan, 2004. [Online]. Available: http://users.isy.liu.se/johanl/yalmip

[16] R. Olfati-Saber and R. Murray, "Consensus problems in networks of agents with switching topology and time-delays," *IEEE Transactions on Automatic Control*, vol. 49, no. 9, pp. 1520–1533, Sept. 2004.

[17] A. Jadbabaie, J. Lin, and A. S. Morse, "Coordination of groups of mobile autonomous agents using nearest neighbor rules," *IEEE Transactions on Automatic Control*, vol. 48, no. 6, pp. 988–1001, 2003.

[18] L. Xiao, S. Boyd, and S. Lall, "A scheme for robust distributed sensor fusion based on average consensus," in *Information Processing in Sensor Networks, 2005. IPSN 2005. Fourth International Symposium on*, april 2005, pp. 63 – 70.

[19] A. Aswani, N. Master, J. Taneja, V. Smith, A. Krioukov, D. Culler, and C. Tomlin, "Identifying models of HVAC systems using semi-parametric regression," in *In Proceedings of the American Control Conference*, 2012.