Using Game Theory for Distributed Control Engineering

Anders Rantzer

Automatic Control LTH
Lund University
July 2008
Abstract

The purpose of this paper is to show how ideas from game theory and economics may play an important role for decentralized controller design in complex engineering systems. The focus is on coordination through prices and iterative price adjustments using a gradient method. We give a quantitative bound on adjustment rates that is sufficient to guarantee global convergence towards a Nash equilibrium of control strategies. The equilibrium is an optimal solution to the corresponding team decision problem, where a distributed set of controllers cooperate to optimize a common objective. The method is illustrated on control of a vehicle formation (e.g. automobiles on the road) where the objective is to maintain desired vehicle distances in presence of disturbances.
Abstract—The purpose of this paper is to show how ideas from game theory and economics may play an important role for decentralized controller design in complex engineering systems. The focus is on coordination through prices and iterative price adjustments using a gradient method. We give a quantitative bound on adjustment rates that is sufficient to guarantee global convergence towards a Nash equilibrium of control strategies. The equilibrium is an optimal solution to the corresponding team decision problem, where a distributed set of controllers cooperate to optimize a common objective. The method is illustrated on control of a vehicle formation (e.g. automobiles on the road) where the objective is to maintain desired vehicle distances in presence of disturbances.

I. INTRODUCTION

How should control equipments distributed across the electricity network cooperate to find new transmission routes when a power line is broken? How should the electronic stabilization programme of an automobile use measurements from wheels and suspensions and decide how to use available brakes and engine power to recover from a dangerous situation? How should radio transmission power between cellphones and base station in a telephone network be adjusted to accommodate for an optimal use of the radio channel when the network load is high? These are all problems of distributed control engineering, where several units need to cooperate with access to different information and with bounds on the communication between them. The classical engineering approach to such problems is to assign one controller for each task and minimize the interaction between them. However, the increasing complexity of engineering systems makes it desirable to go beyond the traditional methods and create a systematic theory for decentralized decision-making and policy updates in dynamical systems. The purpose of this paper is to show how ideas from game theory and economics may play an important role for this purpose.

We consider systems described by differential equations or difference equations. There is a set of agents, each equipped with some decision variables influencing the dynamics of the system. Every agent tries to optimize his own objective defined in terms of the system dynamics. However, the decisions by one agent will also influence the others and we seek methods to handle this interaction through iterative negotiations between the agents.

Iterative processes with provable convergence to a Nash equilibrium for general classes of games are hard to obtain. Similar difficulties appear in general equilibrium theory of economics when it comes to price negotiations aiming to reach a Walras equilibrium. However, many engineering applications can be viewed as “team decision problems”, with an over-all design objective that is common to all agents [6]. This is a special class of games, which are considerably easier to analyze. Introduction of prices makes it possible to split the team problem into an equivalent game where each agent has a local objective which is “linear in money”. A classical argument then shows that price iterations in the gradient direction converge towards the desired equilibrium. See [12, Example 7, page 105].

For a convex-concave function, gradient dynamics in continuous time were proved by Arrow, Hurwicz and Usawa to converge globally towards the saddle-point [1]. The gradient iteration is known as the saddle point algorithm, or Usawa's algorithm. The method to decompose a team problem has been used extensively in methods for large-scale optimization, where it is known as dual decomposition.

Distributed control problems and the relationship to team decision problems have recently gained renewed attention in the engineering literature. It has been shown that a collection of controllers with access to different sets of measurements can be designed using finite-dimensional convex optimization to act optimally as a team. The study of dynamic team problems was initiated already in 1968 by Witsenhausen [13], who also pointed out a fundamental difficulty related to information propagation. Some special types of team problems were solved in the 1970's [11], [5], but the research activity in the area remained moderate until recently. Distributed control problems with spatial invariance was exploited in [2], [3] and conditions for convexity were derived in [10], [9].

In our previous paper [7] a linear quadratic stochastic optimal control problem was solved for a state feedback control law with covariance constraints. The method gives a non-conservative extension of linear quadratic control theory to distributed problems with bounds on the rate of information propagation. An output feedback version of the problem was solved in [8] and for both finite and infinite time horizons in [4].
Finally, we return to the case study in section VI. an informal discussion of vehicle formation control. We introduce some main ideas of the paper through a vehicle system. Assume that the first vehicle controls its own cost function paid for the influence on vehicle 2. Similarly, let x_{11} denote the desired position of vehicle 1. The second vehicle controls x_{22} and pays a cost that depends on x_{33} and x_{22}. Our objective is to find a distributed scheme for coordination of the vehicle movements.

For this purpose, we will use a price mechanism that puts an additional cost on vehicle 1 for every movement of x_{11} that makes life harder for vehicle 2. The purpose is to find prices that create consensus among the vehicles about their desirable positions. Let x_{12} denote the opinion of vehicle 2 regarding the desired position of vehicle 1. Similarly, let x_{23} the position of vehicle 2 as requested by vehicle 3.

To coordinate the movements, prices p_1 and p_2 will be associated with the deviations x_{11} - x_{12} and x_{22} - x_{23} respectively. In particular, the first vehicle will be controlled to minimize the cost function

\[ V_1(x_{11} - v) + p_1 x_{11} \]

Here, the first term represents the cost for first vehicle due to deviations in x_{11} from the desired position v. The second term p_1 x_{11} is the price to be paid for the influence on vehicle 2.

Similarly, the second vehicle is concerned with the cost function

\[ V_2(x_{12}, x_{22}) - p_1 x_{12} + p_2 x_{22} \]

Here the first term is the direct cost that would be generated with the first vehicle positioned in x_{12}. The remaining terms are penalties (or rewards) for the influence on neighboring vehicles. Finally, the third vehicle selects x_{23} and x_{33} to minimize the cost

\[ V_3(x_{23}, x_{33}) - p_2 x_{23} \]  

Let us now consider what happens when for given values of (v, p_1, p_2) each vehicle tries to minimize its cost by adjusting the quantities x_{ij} in the direction of the gradient. This gives the continuous time dynamics

\[ \dot{x}_{11} = -g_{11}[\nabla V_1(x_{11} - v) + p_1] \]  

\[ \dot{x}_{12} = -g_{12}[\nabla V_2(x_{12}, x_{22}) - p_1] \]

\[ \dot{x}_{22} = -g_{22}[\nabla V_2(x_{12}, x_{22}) + p_2] \]

\[ \dot{x}_{23} = -g_{23}[\nabla V_3(x_{23}, x_{33}) - p_2] \]

\[ \dot{x}_{33} = -g_{33}[\nabla V_3(x_{23}, x_{33})] \]

where the numbers g_{11}, g_{12}, g_{22}, g_{23} > 0 specify the rates of adjustment as functions of the cost gradients. At the same time, it is natural to adjust the prices to counteract the deviations x_{11} - x_{12} and x_{22} - x_{23}:

\[ p_1 = h_1(x_{12} - x_{11}) \]  

\[ p_2 = h_2(x_{23} - x_{22}) \]

with rates given by h_1, h_2 > 0. An interesting interpretation of these update rules is that between every pair of neighboring vehicles there is an independent agent trying to benefit from potential disagreements between the two neighbors. In this case, the two intermediate “market makers” would try to adjust their prices to minimize the expressions

\[ p_1(x_{12} - x_{11}) \]

\[ p_2(x_{23} - x_{22}) \]

by adjusting prices in the appropriate direction.

It is a striking fact that, regardless of adjustment rates, under the condition that V_1, V_2 and V_3 are all convex and v is constant, the gradient dynamics turn out to be globally converging towards an equilibrium (x_{11}, x_{12}, x_{22}, x_{23}, x_{33}, p_1, p_2) identical to the saddle point

\[ \max_{p_1, p_2} \min_{x_{ij}} \left[ V_1(x_{11} - v) + V_2(x_{12}, x_{22}) + V_3(x_{23}, x_{33}) + p_1(x_{11} - x_{12}) + p_2(x_{22} - x_{23}) \right] \]

The limiting point is a Nash equilibrium for the game of five players (three vehicles and two market makers) in the sense that no player can reduce his cost by modifying only his own decision variable. Moreover, the equilibrium (x_{11}^*, x_{12}^*, x_{22}^*, x_{23}^*, x_{33}^*) is independent of the adjustment rates and solves the problem

Minimize \[ V_1(x_{11} - v) + V_2(x_{12}, x_{22}) + V_3(x_{23}, x_{33}) \]

subject to \[ x_{11} = x_{12}, \quad x_{22} = x_{23} \]
At the same time, \( (p_1^0, p_2^0) \) are the Lagrange multipliers corresponding to the two equality constraints.

Below is a formal statement of the convergence result (without the sign constraints on \( x_j \) of the original paper) [1]:

**Theorem 1 (Arrow, Hurwicz, Usawa):** Assume that \( V \in C^1(\mathbb{R}^n) \) is strictly convex with gradient \( \nabla V \), while \( G \) and \( H \) are positive definite and \( R \) has full row rank. Then, all solutions to

\[
\dot{x} = -G[(\nabla V)^T - R^T p] \\
p = -HRx
\]

converge to the unique saddle point \((x_*, p_*)\) attaining

\[
\max_{x} \min_{p} \left[ V(x) - p^T Rx \right]
\]

(5)

**Proof.** Let \( \phi(x, p) = V(x) - p^T Rx \). Then

\[
\dot{x} = G[\nabla_x \phi(x, p)]^T \\
p = -H[\nabla_p \phi(x, p)]^T
\]

Define the Lyapunov function

\[
W(x, p) = \frac{1}{2} \left( |x - x_*|^2_{G^{-1}} + |p - p_*|^2_{H^{-1}} \right)
\]

Then

\[
W = \dot{x}^T G^{-1}(x - x_*) + p^T H^{-1}(p - p_*) \\
= [\nabla_x \phi(x, p)](x - x_*) - [\nabla_p \phi(x, p)](p - p_*) \\
\leq [\phi(x, p) - \phi(x_*, p_*]) - [\phi(x, p) - \phi(x_*, p)] \\
= [\phi(x, p) - \phi(x_*, p_*]) - [\phi(x_*, p) - \phi(x_*, p)] \leq 0
\]

with equality if and only if \( x = x_* \). Hence, by LaSalle’s theorem, \((x(t), p(t))\) tends towards \( M \), the largest invariant set in the subspace \( x = x_* \). Invariance means that \( \dot{x} = 0 \), hence \( \nabla V(x^T) = R^T p \), so the only point in \( M \) is \((x_*, p_*)\). This completes the proof. \( \square \)

To concretize the result for the vehicle formation, the response to a brief disturbance in \( v \) is plotted in Figure 2, using gradient dynamics when (1)-(3) are given by

\[
6(x_{11} - v)^2 + p_1 x_{11} \\
3(x_{12} - x_{22})^2 - p_1 x_{12} + p_2 x_{22} \\
2(x_{23} - x_{33})^2 + 2(x_{33})^2 - p_2 x_{23}
\]

Notice that the cost of the first vehicle quickly recovers, but there is a poorly damped oscillation in the response of the second and third vehicle. This reflects the fact that only the stationary equilibrium is optimized, not the transient dynamics. Hence it is natural to ask: Can the same idea of dual decomposition be used to get a distributed scheme for design of dynamic controllers? To address this question, a more abstract version of the theory will be introduced in the next few sections, before returning to the vehicle formation problem.

III. Basic notions of game theory

A (strategic) game is defined by a map

\[
\mu = (\mu_1, \ldots, \mu_J) \mapsto (V_1(\mu), \ldots, V_J(\mu))
\]

where \( \mu_j \) is the strategy of player \( j \) and \( V_j(\mu) \in \mathbb{R} \) is the payoff for player \( j \). The notation \((\mu_1, \tilde{\mu}_{-j})\) is used to denote set of strategies that is equal to \( \mu \) in all entries except \( \mu_j \). A Nash equilibrium of the game is a set of strategies \( \tilde{\mu} = (\tilde{\mu}_1, \ldots, \tilde{\mu}_J) \) such that

\[
V_j(\tilde{\mu}) \geq V_j(\mu_j, \tilde{\mu}_{-j}) \text{ for all } \mu_j
\]

Given \( \epsilon > 0 \), a Nash \( \epsilon \)-equilibrium of the game is a set of strategies \( \tilde{\mu} = (\tilde{\mu}_1, \ldots, \tilde{\mu}_J) \) such that

\[
V_j(\tilde{\mu}) + \epsilon \geq V_j(\mu_j, \tilde{\mu}_{-j}) \text{ for all } \mu_j
\]

A potential game is a game for which there exists a potential function \( \Phi : \mu \mapsto \mathbb{R} \) such that

\[
\text{sgn} [V_j(\tilde{\mu}) - V_j(\mu)] = \text{sgn} [\Phi(\tilde{\mu}) - \Phi(\mu)]
\]

whenever \( \mu_i = \tilde{\mu}_i \) for \( i \neq j \).

For the purposes of this paper, we also introduce the following slightly more general concept: A minmax potential game is a game for which there exists a potential function \( \Phi : \mu \mapsto \mathbb{R} \) and a partitioning \( \{1, \ldots, J\} = \mathcal{F}_1 \cup \mathcal{F}_2 \) such that

\[
\text{sgn} [V_j(\tilde{\mu}) - V_j(\mu)] = \begin{cases} 
\text{sgn} [\Phi(\tilde{\mu}) - \Phi(\mu)] & \text{if } j \in \mathcal{F}_1 \\
-\text{sgn} [\Phi(\tilde{\mu}) - \Phi(\mu)] & \text{if } j \in \mathcal{F}_2
\end{cases}
\]

whenever \( \mu_i = \tilde{\mu}_i \) for \( i \neq j \). The players in \( \mathcal{F}_1 \) maximize the potential, while the players in \( \mathcal{F}_2 \) minimize the potential.

A team problem with team payoff \( V' \) is a game where \( V'_1 = \ldots = V'_J = V' \). If \( V'(\tilde{\mu}) = \max_\mu V'(\mu) \), then \( \tilde{\mu} \) is called an optimal set of strategies for the team problem. (Moreover, \( \tilde{\mu} \) is a Nash equilibrium of every potential game with potential function \( V' \).)
IV. DISTRIBUTED TEAM PROBLEMS

We will now study how a team problem with graph structure can be decomposed into subproblems with individual objectives. Given a graph with nodes \( J = \{1,2,\ldots,J\} \) and a set of edges \( E \), consider a team problem with strategies \( \mu_1, \ldots, \mu_J \) associated with the nodes. Consider first a team problem with payoff of the form

\[
\mathcal{V}(\mu) = V_1(\mu_{[1]}) + \cdots + V_J(\mu_{[J]})
\]

where \( \mu_{[j]} \) denotes the collection of strategies in node \( j \) and neighboring nodes \( i \) such that \((j,i) \in E\). The problem can be decomposed and a distributed iteration \((\hat{\mu},\hat{\tau})\) can be achieved as in the following theorem, where the notation

\[
\mu_j(\tau + 1) = \arg\max_{\mu_j} \mathcal{V}_j(\mu_j, \mu_{-j}(\tau))
\]

represents the map \( \mu(\tau) \rightarrow \mu_j(\tau + 1) \) defined by

\[
\mu_j(\tau + 1) = \begin{cases} 
\arg\max_{\mu_j} \mathcal{V}_j(\mu_j, \mu_{-j}(\tau)) & \text{if } \max_{\mu_j} \mathcal{V}_j(\mu_j, \mu_{-j}(\tau)) > \mathcal{V}_j(\mu_j(\tau)) + \epsilon \\
\mu_j(\tau) & \text{otherwise.}
\end{cases}
\]

Theorem 2: Let

\[
\mathcal{V}_j(\mu) = \sum_{(i,j) \in E} V_i(\mu_{[i]})
\]

Then the game defined by (6) is a potential game with the potential \( \mathcal{V}(\mu) \). A Nash equilibrium for this potential game is a Nash equilibrium for the team problem with payoff \( \mathcal{V}(\mu) \) and vice versa.

Moreover, suppose that \( \mathcal{V}(\mu) \) is bounded from above. Then, a strategy sequence \( (0, 1, 2, \ldots) \) reaches a Nash \( \epsilon \)-equilibrium of the potential game after a finite number of iterations, provided that the sequence \( j_1, j_2, \ldots \) visits every node sufficiently many times.

Remark 1. Unlike the team problem, the payoffs of in the potential game are local, in the sense that \( \mathcal{V}_j(\mu) \) is entirely determined by the strategies in node \( j \) and neighboring nodes. Hence, the iterative updates are also local in the sense that each node only needs to exchange information with neighboring nodes in the graph.

Proof. When \( \mu_i = \hat{\mu}_i \) for \( i \neq j \), the definitions of \( \mathcal{V} \) and \( \mathcal{V}_j \) give

\[
\mathcal{V}_j(\hat{\mu}) - \mathcal{V}_j(\mu) = \mathcal{V}(\hat{\mu}) - \mathcal{V}(\mu)
\]

Hence the game defined by (6) is a potential game with the potential \( \mathcal{V}(\mu) \). The equality also shows that a Nash equilibrium for the potential game is a Nash equilibrium for the team problem with payoff \( \mathcal{V}(\mu) \) and vice versa.

The construction of the sequence \( \{\mu(\tau)\}_{\tau=0}^{\infty} \) gives

\[
\mathcal{V}(\mu(0)) \leq \mathcal{V}(\mu(1)) \leq \mathcal{V}(\mu(2)) \leq \cdots
\]

The sequence is bounded from above, so there exists \( T > 0 \) with

\[
\mathcal{V}(\mu(T)) = \lim_{\tau \to \infty} \mathcal{V}(\mu(\tau)) - \epsilon
\]

The iteration (10) shows that \( \mu(\tau) \) remains constant for \( \tau \geq T \) and because every node is visited by \( j \) for some \( \tau \geq T \), the final strategy \( \mu(T) \) must be a Nash \( \epsilon \)-equilibrium.

V. STRATEGIES WITH GLOBAL IMPACT

For team payoffs of the form (9), modifying the strategy in a single node will only influence the payoff in neighboring nodes. A more advanced team problem, where this is no longer the case, will be introduced next.

Let \((i_1, j_1), (i_2, j_2), \ldots, (i_K, j_K)\) be an enumeration of the graph edges, i.e. the elements of \( E \). Define the team payoff

\[
\mathcal{V}(\mu) = V_1(x_1, \mu_1) + \cdots + V_J(x_J, \mu_J)
\]

where the vector \( x = (x_1, \ldots, x_J) \) is supposed to be uniquely determined from \( \mu_1, \ldots, \mu_J \) by the equations

\[
0 = a_k(x_{i_k}) + b_k(x_{j_k}, \mu_{j_k}) \quad k = 1, \ldots, K
\]

Interestingly, this game can be decomposed by introduction of an appropriate minmax potential game. The potential game has strategies \( \{\mu_1, x_1, \ldots, \mu_J, x_J\} \) corresponding to the nodes and \( \lambda_1, \ldots, \lambda_K \) corresponding to the edges. It is assumed that \( \lambda_k \) belongs to a Hilbert space where also \( a_k \) and \( b_k \) take their values. The Hilbert space inner product is denoted \( \langle \cdot, \cdot \rangle \). Define node payoffs

\[
\mathcal{V}_j(\mu, x, \lambda) = V_j(x_j, \mu_j) + \sum_{i_j = j} \langle \lambda_j, a_k(x_{i_j}) \rangle + \sum_{i_j \neq j} \langle \lambda_j, b_k(x_{i_j}, \mu_{i_j}) \rangle
\]

and edge payoffs

\[
W_k(\mu, x, \lambda) = -\langle \lambda_k, a_k(x_{i_k}) + b_k(x_{j_k}, \mu_{j_k}) \rangle
\]

Consider the game with node strategies \( \{\mu_1, x_1\}, \ldots, \{\mu_J, x_J\} \), edge strategies \( \lambda_1, \ldots, \lambda_K \) and payoffs

\[
\mathcal{V}_1(\mu, x, \lambda), \ldots, \mathcal{V}_J(\mu, x, \lambda), W_1(\mu, x, \lambda), \ldots, W_K(\mu, x, \lambda)
\]

This is a minmax potential game with potential function

\[
\Phi(\mu, x, \lambda) = \sum_{j=1}^{J} \mathcal{V}_j(x_j, \mu_j) + \sum_{k=1}^{K} \langle \lambda_k, a_k(x_{i_k}) + b_k(x_{j_k}, \mu_{j_k}) \rangle
\]

The node players maximize the potential, while the edge players minimize it. Then the following theorem holds.
Theorem 3: Given $\mathcal{V}_1, \ldots, \mathcal{V}_q, W_1, \ldots, W_k$ as in (13)-(14), the corresponding game is a minmax potential game with potential $\Phi(\mu, x, \lambda)$ defined by (15). If $(\tilde{\mu}, \tilde{x}, \tilde{\lambda})$ is a Nash equilibrium of this game, then $\tilde{\mu}$ is a Nash equilibrium for the team problem defined by (11)-(12).

Moreover, suppose $\min_{\lambda} \max_{\mu, x} \Phi(\mu, x, \lambda)$ is attained in a point $(\tilde{\mu}, \tilde{x}, \tilde{\lambda})$ satisfying the following (local) conditions: There exists $\varepsilon > 0$ such that for $\lambda$ in a neighborhood of $\tilde{\lambda}$

$$\max_{\mu, x} \Phi(\mu, x, \lambda) \geq \Phi(\tilde{\mu}, \tilde{x}, \tilde{\lambda}) + \varepsilon \|\lambda - \tilde{\lambda}\|^2$$

and

$$\arg \max_{\mu, x} \Phi(\mu, x, \lambda) = (\tilde{\mu}, \tilde{x}, \tilde{\lambda})$$

as $\lambda \to \tilde{\lambda}$. Then, the sequence $\{(\mu(t), x(t), \lambda(t))\}_{t=1}^{\infty}$ defined by

$$(\mu(0), x(0), \lambda(0)) = (\tilde{\mu}, \tilde{x}, \tilde{\lambda}) \quad \text{and} \quad (\mu(t+1), x(t+1), \lambda(t+1)) = \arg \max_{(\mu, x)} \left( \mathcal{V}(x(t), \mu(t)) + \sum_{i=1}^{k} \langle \lambda_i, a_i(x_i(t)) \rangle + \sum_{j=1}^{n} \langle \lambda_j, b_j(x_j(t), \mu_j(t)) \rangle \right)$$

converges (globally) towards the Nash equilibrium $(\tilde{\mu}, \tilde{x}, \tilde{\lambda})$ as $\tau \to \infty$, provided that the numbers $\gamma_1, \ldots, \gamma_k > 0$ are small enough to make $\max_{\mu, x} \Phi(\mu, x, \lambda) - \sum_{k} \|\lambda_k\|^2 / \eta_k$ a concave function of $\lambda$ for some $\eta_k > 0$.

Proof. When $(\mu(t), x(t)) = (\tilde{\mu}, \tilde{x})$ for $t \neq 0$, we have

$$\mathcal{V}(\tilde{\mu}, \tilde{x}, \lambda) - \mathcal{V}(\mu(t), x(t)) = \Phi(\tilde{\mu}, \tilde{x}, \lambda) - \Phi(\mu(t), x(t), \lambda)$$

Similarly, when $\lambda(t) = \tilde{\lambda}$ for $t \neq k$

$$W_k(\tilde{\mu}, \tilde{x}, \lambda) - W_k(\mu(t), x(t), \lambda) = \Phi(\mu(t), x(t), \lambda) - \Phi(\tilde{\mu}, \tilde{x}, \tilde{\lambda})$$

Hence the conditions for a minmax potential game hold. If $(\tilde{\mu}, \tilde{x}, \tilde{\lambda})$ is a Nash equilibrium of this game, then

$$\mathcal{V}(\mu) \leq \sup_{(\mu, \lambda)} \inf_{x} \Phi(\mu, x, \lambda) \leq \sup_{(\mu, \lambda)} \Phi(\tilde{\mu}, \tilde{x}, \tilde{\lambda}) = \mathcal{V}(\tilde{\mu})$$

The first inequality follows directly from the definition of $\mathcal{V}(\mu)$, while the two equalities follow from the assumption that $(\tilde{\mu}, \tilde{x}, \tilde{\lambda})$ is a Nash equilibrium. Altogether, this proves that $\tilde{\mu}$ is a Nash equilibrium for the team problem.

To prove convergence towards the Nash equilibrium, introduce

$$h = \begin{bmatrix} h_1 \\ \vdots \\ h_k \end{bmatrix}$$

$$\langle \lambda, \eta^{-1} \lambda \rangle = \sum_{k} \langle \lambda_k, \lambda_k \rangle / \eta_k$$

We also use the notation $\rho = \lambda(\tau + 1), x^+ = x(\tau + 1)$ and $\mu^+ = \mu(\tau + 1)$. Let $U(\lambda) = \max_{\mu, x} \Phi(\mu, x, \lambda)$. Then

$$U(\lambda) - U(\tilde{\lambda}) = \Phi(\mu^+, x^+, \lambda) - \max_{\mu, x} \Phi(\mu, x, \lambda)$$

$$\leq \Phi(\mu^+, x^+, \lambda) - \Phi(\mu^+, x^+ + h, \lambda)$$

$$= \langle \lambda - \tilde{\lambda}, a(x^+) + b(x^+, \mu^+) \rangle$$

$$= \langle \lambda - \tilde{\lambda}, 2 \eta^{-1} h \rangle$$

The condition (16) can be written

$$U(\lambda) \geq U(\tilde{\lambda}) + \varepsilon \|\lambda - \tilde{\lambda}\|^2$$

for $\lambda$ close to $\tilde{\lambda}$. However, $U$ is convex by construction, so

$$U(\lambda) \geq U(\tilde{\lambda}) + \|\lambda - \tilde{\lambda}\| \min_{\mu} \{-\langle \lambda - \tilde{\lambda}, \mu \rangle\}$$

must hold globally for sufficiently small $\mu > 0$. Combining this with (17) gives

$$\|\lambda - \tilde{\lambda}\| \min_{\mu} \{-\langle \lambda - \tilde{\lambda}, \mu \rangle\} \leq \langle \lambda - \tilde{\lambda}, 2 \eta^{-1} h \rangle$$

Thus

$$\min_{\mu} \{-\langle \lambda - \tilde{\lambda}, \mu \rangle\} \leq 2 \|\mu\|^2 h$$

To prove that $\lim_{\tau \to \infty} \langle \lambda(\tau), \mu(\tau) \rangle = 0$, it therefore remains to prove that $\lim_{\tau \to \infty} \|h(\tau)\| = 0$. Concavity of $U(\lambda) - \langle \lambda, \eta^{-1} \lambda \rangle$ gives

$$U(\lambda + h) - \langle \lambda + h, \eta^{-1} (\lambda + h) \rangle$$

$$= \Phi(\lambda) - \langle h, a(x^+) \rangle + \Phi(\mu^+) \leq 2 \langle \eta^{-1} h, \|h\|^2 \rangle$$

In particular

$$U(\lambda) = U(\lambda(T)) - \sum_{t=0}^{T-1} \left( U(\lambda(t+1)) - U(\lambda(t)) \right) + U(\lambda(0))$$

$$\leq 2(\eta^{-1} h) \sum_{t=0}^{T-1} \|h(t)\|^2 + U(\lambda(0))$$

so $\sum_{t=0}^{T} \|h(t)\|^2$ is bounded and $\lim_{\tau \to \infty} \|h(\tau)\| = 0$. This proves that $\lim_{\tau \to \infty} \lambda(\tau) = \tilde{\lambda}$. By assumption, this also gives $\lim_{\tau \to \infty} \mu(\tau, x(\tau)) = (\tilde{\mu}, \tilde{x})$, so the proof is complete. \ \square

VI. Vehicle formations reconsidered

For optimization of dynamic controllers in the vehicle formation, it is useful to quantify the resulting stationary dynamics when $v$ is a given discrete time stochastic process. For an example of stable (but sub-optimal) dynamics, one could let each vehicle optimize its position based on prices determined by the following discrete time version of the gradient dynamics studied before:

$$p_1(t+1) = p_1(t) + h_1 (x_{11}(t) - x_{12}(t))$$

$$p_2(t+1) = p_2(t) + h_2 (x_{22}(t) - x_{23}(t))$$

(18)

For small adjustment rates, the discrete time dynamics is similar to the continuous time behavior, hence stable. Suppose the disturbance $w(t)$ is a Gaussian discrete time stationary stochastic process given by

$$v(t+1) = av(t) + w(t)$$

where $w(t)$ is zero mean white noise with unit variance. Figure 3 shows a simulation with $a = 0.9$ and $g_{11} = g_{12} = g_{22} = g_{23} = g_{33} = h_1 = h_2 = 0.1$. Due to the structure of the gradient dynamics, the effect of $w(t)$ will propagate through the vehicle formation, affecting the first vehicle position $x_{11}$ at time $t+1$, the price $p_1$ and the second vehicle position at time $t+2$ and finally the price $p_2$ and the third vehicle at time $t+3$. We will now consider iterative improvement of
the prices subject to the same information constraints, i.e. $x_{11}(t + 1)$ should be determined based on knowledge of $u$ until time $t$, etc.

Suppose that the vehicles are initially run for some time using fixed fixed price strategies $\lambda_1, \lambda_2$, where $\lambda_1$ maps past values of $x_{12}$ and $x_{11}$ into the current value of $p_1$ and $\lambda_2$ maps past values of $x_{22}$ and $x_{21}$ into the current value of $p_2$. Each market maker collects data during this period and uses this data to study the time correlations between his prices and the local vehicle positions. The measurements make it possible for him to adjust his price strategies for the next time period according to Theorem 3. Once all strategies have been updated, we let the system run again with the new price strategies. The improvement is significant, as shown in Figure 4.

Iterating this procedure over and over again gives a sequence of policies that (as predicted by Theorem 3) converges towards a Nash equilibrium that minimizes the sum of the expected costs in the three vehicles. A simulation with the price optimal policies shows that the previous variations eventually disappear almost entirely. The synthesis procedure is completely distributed and a global optimum is reached even though no agent has a complete model of the entire vehicle formation.

VII. CONCLUDING REMARKS

We have shown that the classical idea of dual decomposition can be applied to a general class of team decision problems. This opens up a wide range of potential applications in distributed control engineering and economics.

VIII. ACKNOWLEDGMENT

The work was supported by a Senior Individual Grant of Swedish Foundation for Strategic Research.

REFERENCES