A Distributed Linear-Quadratic Discrete-Time Game Approach to Multi-Agent Consensus

Prima Aditya and Herbert Werner

Abstract— **This article considers a distributed game theoretic approach to the multi-agent consensus problem, where we represent the consensus problem with multiple decision-makers as a Linear Quadratic Discrete-Time Game (LQDTG) over a connected communication graph. The players - taken here as double integrators - minimize individual finite-horizon cost functions and aim at reaching a Nash equilibrium. This involves a set of coupled Riccati difference equations, which cannot be solved in a distributed manner. Taking up an idea reported previously in the literature in a continuous-time setting, we propose a distributed strategy for reaching a Nash equilibrium that is based on an associated multi-agent system evolving not on the vertices but on the edges of the underlying communication graph. This makes it possible to move the coupling between agents from the cost function to the agent dynamics, and to formulate an LQR problem with a decoupled cost that allows a distributed solution on the edges of the graph. Mapping this solution to local control inputs for the real agents (the nodes of the graph) requires however knowledge of the whole network. To arrive at a distributed control strategy, we propose an iterative, gradient-based construction of the individual control inputs that is based only on locally available information, but requires repeated exchange of information between neighboring agents within a single sampling interval. We show that with a suitable choice of terminal cost weights the distributed solution to the LQR problem is identical to the solution leading to the original Nash equilibrium. The resulting distributed strategy can be implemented in a receding horizon manner, and we provide a sufficient stability condition (assuming convergence of the iterative input construction) in terms of the terminal cost weight.**

I. INTRODUCTION

The multi-agent consensus control problem has been widely studied as an essential application of cooperative control theory. It requires several agents to agree on a common reference, and is considered a fundamental issue in controlling multi-agent systems [1]. It is also the basis for formation control, where offsets are introduced to make agents converge to a predetermined geometric shape. To attain and preserve consensus, agents need to exchange information such as position and velocity with their neighbors.

A multi-agent consensus problem can be modeled as a game where each agent in the game is pursuing its individual strategy, while exchanging information locally with its neighbors. This game scenario precisely fits the differential game formulation: instead of one agent pursuing one goal, there are several agents, all pursuing their individual goals which may or may not be in conflict with other agents' goals.

Prima Aditya and Herbert Werner are with Institute of Control Systems, Hamburg University of Technology, 21073 Hamburg, Germany {prima.aditya, h.werner}@tuhh.de

This paper considers on a non-cooperative game formulation for a homogeneous system of double integrator agents: each agent minimizes its objective by considering other players' strategies, leading to a Nash equilibrium where no player in a game can enhance its payoff by unilaterally changing the strategy [2].

Common practice in differential games is to represent the objective function as linear-quadratic, leading to LQ differential games (LQDG) [3]. While commonly studied in the continuous-time domain, discrete-time versions of it have also been proposed; see [4]. The differential game approach to formation based consensus control was introduced in [5]. The task of all agents is to achieve consensus (or attain a formation) as solution of a game, i.e. to find a Nash equilibrium. In this approach, the problem requires a system of coupled Riccati equations to be solved, which can however not be done in a distributed manner and which will become computationally expensive or even intractable when solved centrally and the number of agents is large.

A discrete-time version of this approach was proposed in [6], where agents solve (coupled) Riccati difference equations backward in time to obtain the solution of the game. As it is more practical with a view on a receding horizon implementation, here we focus on LQ discrete-time games (LQDTG).

From a practical point of view, it is desirable to implement a consensus or formation control scheme in a distributed manner, i.e. agents should be able to determine their local strategy based on locally available information. The work presented here is inspired by the approach in [7], where it was shown how to avoid the need to solve coupled (asymmetric) Riccati equations, by relocating the coupling terms from the cost function to relative dynamics. This is achieved by considering a fictitious multiagent system that is evolving on the edges rather than on the vertices of the interaction graph (which in [7] is restricted to a line topology). As a result, for the new problem one can solve a family of decoupled (symmetric) Riccati equations independently. A problem not addressed in [7] is that mapping the optimal control sequence on the edges to the one on the nodes (i.e. the control input to the real agents) still requires information across the whole network. Here we propose a distributed approach that is based on solving the associated linear system of equations iteratively, exploiting the fact that the gradient of the error depends only on locally available information.

Contributions in this article are:

1. An extension of the approach in [7] to general interaction graphs (where we assume the number of edges

is larger than the number of vertices), and to discrete-time games
2. Use of a distributed steepest descent algorithm that allows to calculate the local control inputs based only on locally available information
3. Establishment - after appropriately scaling the terminal cost - of the equivalence between the decoupled solution on the edge system and the coupled solution to the original problem evolving on the vertices
3. Sufficient stability conditions for a receding horizon implementation of the proposed approach (assuming convergence of the iterative input construction)

II. PRELIMINARIES AND PROBLEM STATEMENT

A. Graph Theory

An directed graph $\mathcal{G} := (\mathcal{V}, \mathcal{E})$ consists of a set of vertices $\mathcal{V} = \{\nu^1, ..., \nu^N\}$, and a set of edges $\mathcal{E} = \{(\nu^i, \nu^j) \in \mathcal{V} \times \mathcal{V}, \nu^j \neq \nu^i\}$, containing ordered pairs of distinct vertices. N is the number of vertices and M the number of edges. We denote an edge by $e^{ij} := (\nu^j, \nu^i)$ and its associated weight by $\mu^{ij} \geq 0$. The set of neighbors of agent i (sharing a common edge) is denoted by $\mathcal{N}^i$. The incidence matrix $D \in \mathbb{R}^{N\times M}$ of a directed graph has an ij element with value 1 if the node i is the head of the edge e^{ij}, -1 if the node i is the tail, and 0 otherwise. When constructing the incidence matrix for an undirected graph, head and tail of an edge can be assigned arbitrarily. We define the weighted Laplacian of a graph $\mathcal{G}$ as

$$L = DWD^T, \tag{1}$$

where $W = \text{diag}(..., \mu^{ij}, ...) \in \mathbb{R}^{M\times M}$. The Laplacian matrix is symmetric and positive semi-definite.

B. Agent Model and Performance Indices

We consider multi-agent systems where N agents - seen as point masses - are moving in n-dimensional space. Agents are thus modeled as double integrators; in the context of game theory the local controller of an agent is interpreted as a decision maker in a discrete time game. The discrete-time model of the double integrator dynamics of agent i is

$$x^i_{k+1} = f^i x^i_k + g^i u^i_k, \tag{2}$$

where $x^i_k = [p^{i^T}_k, v^{i^T}_k]^T \in \mathbb{R}^{2n}$ is the state vector that includes position p^i and velocity v^i of agent i, with

$$f^i = \begin{bmatrix} I_n & \delta I_n \\ 0 & I_n \end{bmatrix}, \quad g^i = \begin{bmatrix} \frac{\delta^2}{2} I_n \\ \delta I_n \end{bmatrix},$$

and where u^i_k is the control input (acceleration), while δ denotes the sampling time. We construct a state vector for the whole multi-agent system as $\tilde{x}_k = [x^1_k, ..., x^N_k]^T \in \mathbb{R}^{2Nn}$. To have a consistent representation for the state vector of the whole multi-agent, here we will use a permuted version of the state vector

$$x_k = \Pi_N \tilde{x}_k, \tag{3}$$

where the permutation matrix Π_N is given as

$$\Pi_N = \begin{bmatrix} I_N \otimes \begin{bmatrix} 1 & 0 \end{bmatrix} \\ I_N \otimes \begin{bmatrix} 0 & 1 \end{bmatrix} \end{bmatrix} \otimes I_n \in \mathbb{R}^{2Nn\times 2Nn}. \tag{4}$$

Finally, the overall dynamics of the multi-agent system can be written as

$$x_{k+1} = Fx_k + \sum_{i=1}^N G^i u^i_k, \tag{5}$$

where the permuted state vector now is arranged as $x_k = [p^{1^T}_k, ..., p^{N^T}_k, v^{1^T}_k, ..., v^{N^T}_k]^T \in \mathbb{R}^{2Nn}$, with

$$F = \begin{bmatrix} I_{Nn} & \delta I_{Nn} \\ 0 & I_{Nn} \end{bmatrix} \in \mathbb{R}^{2Nn\times 2Nn}$$

and $G^i = \left(\bar{g}^i \otimes [0, ..., 1, ..., 0]^T\right) \otimes I_n \in \mathbb{R}^{2Nn\times n}$, where $\bar{g}^i = \begin{bmatrix} \frac{\delta^2}{2} & \delta \end{bmatrix}^T$.

We assume that agents are exchanging state information, and that the communication structure is represented by a graph $\mathcal{G}$ that determines for each agent i a set of neighbors $\mathcal{N}^i$. We will make the following two assumptions.

Assumption 1: $\mathcal{G}$ is (weakly) connected.

Assumption 2: We have $M \geq N$, *i.e.* the number of edges is not smaller than the number of vertices.

The objective of a consensus protocol is to make all agents converge to a common reference with respect to (relative) position and velocity (for formation control one can introduce offsets so that agents converge to a specified geometric shape instead of a common point). The local error for agent i to be minimized is defined as

$$\sum_{j\in\mathcal{N}^i} \mu^{ij}\left(||p^i_k - p^j_k||^2 + ||v^i_k - v^j_k||^2\right) = p^T_k \mathbf{L}^i p_k + v^T_k \mathbf{L}^i v_k = x^T_k Q^i x_k, \tag{6}$$

where $Q^i = \delta \begin{bmatrix} \mathbf{L}^i & 0 \\ 0 & \mathbf{L}^i \end{bmatrix} \in \mathbb{R}^{2Nn\times 2Nn}$, $\mathbf{L}^i = L^i \otimes I_n$, with the Laplacian matrix for each agent $L^i = DW^iD^T$, and weights $W^i = \text{diag}(0, ..., \mu^{ij}, ..., 0) \in \mathbb{R}^{M\times M}$, to be diagonal matrix of edge weights, and $j \in \mathcal{N}^i$. The game considered here is assumed to be played in the open-loop finite horizon $[0, T]$ and the cost function for each agent is

$$J^i_k = \frac{1}{2}\Big(X^T_k \bar{Q}^i X_k + \sum_{j=1}^N U^{j^T}_k \bar{R}^{ij} U^j_k\Big), \tag{7}$$

where $X_k = [x^T_{k+1}, x^T_{k+2}, ..., x^T_T]^T$ and $U^i_k = [u^{i^T}_k, u^{i^T}_{k+1} ..., u^{i^T}_{T-1}]^T$. The weighting matrices for the state costs are given by $\bar{Q}^i = \text{blkdiag}(Q^i, ..., Q^i, Q^i_T)$, where $Q^i_T = \delta \begin{bmatrix} \mathbf{L}^i_T & 0 \\ 0 & \mathbf{L}^i_T \end{bmatrix} \in \mathbb{R}^{2Nn\times 2Nn}$, $\mathbf{L}^i_T = L^i_T \otimes I_n$, with $L^i_T = DW^i_T D^T$, and $W^i_T = \text{diag}(0, ..., \eta^{ij}, ..., 0) \in \mathbb{R}^{M\times M}$. The control weight is taken as $\bar{R}^{ij} = 0$ when $i \neq j$, otherwise as $\bar{R}^{ii} = \text{blkdiag}(R^{ii})$ with $R^{ii} \succ 0$. Note that the coupling (exchange of information) between agents is expressed by the presence of the Laplacian in the cost function via Q^i.

C. Nash Equilibrium and Coupled Riccati Difference Equation

The formulation of the consensus (or formation) control problem with dynamics (5) and cost function (7) as a game reflects the non-cooperative agent behavior, where each player minimizing its local cost function leads to a Nash equilibrium.

Definition 1: A collection of strategies $U^{i\star}$ constitutes a Nash equilibrium if and only if the inequalities

$$J^i(U^{1\star}, ..., U^{N\star}) \leq J^i(U^{1\star}, ..., U^{i-1\star}, U^i, U^{i+1\star}, ..., U^{N\star}),$$

hold for $i = 1, ..., N$.

Now we can formulate the consensus problem for the multi-agent system (5) with a given Laplacian matrix L^i and cost function (7) as follows.

Problem 1: Minimize the local cost function (7) over the control input sequences u^i subject to dynamics (5).

Theorem 1: An open-loop Nash equilibrium for the game defined by Problem 1 is achieved by the control sequences

$$u_k^{i\star} = -R^{ii^{-1}} G^{i^T} P_{k+1}^i x_{k+1}, \quad i = 1, ..., N, \tag{8}$$

if P_k^i are solutions to the discrete-time coupled Riccati difference equations

$$P_k^i = Q_k^i + F^T P_{k+1}^i \Big[I + \sum_{j=1}^{N} G^j R^{j^{-1}} G^{j^T} P_{k+1}^j\Big]^{-1} F, \tag{9}$$

which need to be solved backward with $P_T^i = Q_T^i$, where $R^j = R^{jj}$. The corresponding closed-loop state trajectory is

$$x_{k+1} = \Big[I + \sum_{j=1}^{N} G^j R^{j^{-1}} G^{j^T} P_{k+1}^j\Big]^{-1} F x_k. \tag{10}$$

Proof: See [4]. ■

Note that solving for P_k^i requires knowledge of P_{k+1}^j for all $j = 1, \ldots, N$; thus the problem cannot be solved in a distributed manner.

III. A DISTRIBUTED SOLUTION TO LQDTG

In this section we present a distributed solution to the above problem. Taking up an idea developed in [7], we first consider a fictitious multi-agent system, where we associate each edge of the underlying communication graph with an agent that represents the relative dynamics between (vertex) agents i and j. This results in a decoupled cost function and moves the coupling into the agent dynamics, thus leading to a problem that can be solved in a distributed manner. We will then show how to obtain form the the distributed framework and obtain the optimal solution.

A. Consensus in the Edge System

We define the relative dynamics between (vertex) agents i and j as

$$\begin{bmatrix} q_k^m \\ w_k^m \end{bmatrix} = \begin{bmatrix} p_k^i - p_k^j \\ v_k^i - v_k^j \end{bmatrix} \quad \text{and} \quad a_k^m = u_k^i - u_k^j, \tag{11}$$

for $m = 1, ..., M$. The relative dynamics

$$z_{k+1}^m = f^m z_k^m + g^m a_k^m, \qquad \text{for} \qquad m = 1, ..., M, \tag{12}$$

can be associated with a new (edge) agent m, where $z_k^m = [q_k^{m^T}, w_k^{m^T}]^T \in \mathbb{R}^{2n}$ is the state vector and a_k^m is the control input vector of an agent associated with edge e^{ij}. Arrange a relative state vector for the whole system as $\tilde{z}_k = [z_k^1, ..., z_k^M]^T \in \mathbb{R}^{2Mn}$. A permutation yields

$$z_k = \Pi_M \tilde{z}_k, \tag{13}$$

with a permutation matrix

$$\Pi_M = \begin{bmatrix} I_M \otimes \begin{bmatrix} 1 & 0 \end{bmatrix} \\ I_M \otimes \begin{bmatrix} 0 & 1 \end{bmatrix} \end{bmatrix} \otimes I_n \in \mathbb{R}^{2Mn \times 2Mn}.$$

The whole edge system can be written as

$$z_{k+1} = \tilde{F} z_k + \sum_{m=1}^{M} \tilde{G}^m a_k^m, \tag{14}$$

where the permuted state vector is $z_k = [q_k^{1^T}, ..., q_k^{N^T}, w_k^{1^T}, ..., w_k^{N^T}]^T \in \mathbb{R}^{2Mn}$, the system matrix is given by $\tilde{F} = \begin{bmatrix} I_{Mn} & \delta I_{Mn} \\ 0 & I_{Mn} \end{bmatrix} \in \mathbb{R}^{2Mn \times 2Mn}$, and the input matrix is defined as $\tilde{G}^m = \Big(\bar{g}^m \otimes [0, ..., 1, ..., 0]^T\Big) \otimes I_n \in \mathbb{R}^{2Mn \times n}$.

The state error on edge m is

$$\mu^m \Big(||q_k^m||^2 + ||w_k^m||^2\Big) = z_k^T \tilde{Q}^m z_k, \tag{15}$$

where $\tilde{Q}^m = \delta(I_2 \otimes \tilde{W}^m \otimes I_n) \in \mathbb{R}^{2Mn \times 2Mn}$, with $\tilde{W}^m = \text{diag}(0, ..., \tilde{\mu}^m, ..., 0)$ and $\tilde{\mu}^m$ is the mth weight in the set $(..., \tilde{\mu}^{ij}, ...)$ with the same ordering as $\mathcal{E}$. The finite horizon cost function for the mth edge can be expressed as

$$\tilde{J}_k^m = \frac{1}{2}\Big(Z_k^T \hat{Q}^m Z_k + \sum_{j=1}^{M} A_k^{j^T} \hat{R}^{mj} A_k^j\Big), \tag{16}$$

where the relative state vector now is chosen as $Z_k = [z_{k+1}^T, z_{k+2}^T, ..., z_{k+T}^T]^T$, and the relative control input vector is $A_k^m = [a_k^{m^T}, a_{k+1}^{m^T}, ..., a_{k+T-1}^{m^T}]^T$. The state cost weighting is $\hat{Q}^m = \text{blkdiag}(\tilde{Q}^m, ..., \tilde{Q}^m, \tilde{Q}_T^m) \in \mathbb{R}^{2MnT \times 2MnT}$, where $\tilde{Q}_T^m = \delta(I_2 \otimes \tilde{W}_T^m \otimes I_n) \in \mathbb{R}^{2Mn \times 2Mn}$, with $\tilde{W}_T^m = \text{diag}(0, ..., \tilde{\eta}^m, ..., 0)$ and $\tilde{\eta}^m > 0$, is the mth weight in the set $(..., \tilde{\eta}^{ij}, ...)$ with the same ordering as $\mathcal{E}$. The control weight is $\hat{R}^{mj} = \text{blkdiag}(\tilde{R}^{mj})$ with $\tilde{R}^{jj} \succ 0$. For this new system of edge dynamics (14), we can formulate the following problem.

Problem 2: Minimize the cost function (16) over the relative acceleration control input sequences a^i subject to dynamics (14).

Theorem 2: The optimal solution to Problem 2 is

$$a_k^{m\star} = \tilde{K}_k^m z_k, \quad \text{for} \quad m = 1, ..., M \tag{17}$$

where

$$\tilde{K}_k^m = -\Big(\tilde{R}^{mm} + \tilde{G}^{m^T} \tilde{P}_{k+1}^m \tilde{G}^m\Big)^{-1} \tilde{G}^{m^T} \tilde{P}_{k+1}^m \tilde{F}, \tag{18}$$

and $\tilde{P}_k^m$ is the solution to the Riccati difference equation

$$\tilde{P}_k^m = \tilde{Q}^m + \tilde{F}^T \tilde{P}_{k+1}^m \tilde{F} + \tilde{F}^T \tilde{P}_{k+1}^m \tilde{G}^m \tilde{K}_k^m, \quad \tilde{P}_T^m = \tilde{Q}_T^m. \tag{19}$$

Proof: Standard, see e.g. [8]. ∎

Lemma 1: A (vertex) multi-agent system in the form of (5) reaches consensus $\bar{x}$ if the state of the corresponding edge system in (14) approaches zero, i.e.,

$$\lim_{k\to\infty} z_k = 0. \tag{20}$$

Proof: Note that the relationship between vertex dynamics (5) and edge dynamics (14) is

$$z_k = \Phi_z x_k,$$

where $\Phi_z = (I_2 \otimes (-D^T) \otimes I_n) \in \mathbb{R}^{2Mn\times 2Nn}$. By Assumption 1 we have $\operatorname{rank}(\Phi_z) = N-1$ and thus the null space of Φ_z is the agreement space [9]. ∎

We can observe the solution of Problem 2 here that the state feedback $\tilde{K}^m$ is sparse and decoupled from each other – that's why it is making it distributed and can solve the optimization offline and store it. Nevertheless, the direct back transformation from the relative control inputs $a_k^{m\star}$ to the actual control inputs $u_k^{i\star}$ has to be solved in a distributed manner.

B. Distributed Framework Solution

Next, we show how to obtain the control inputs values of real agents back from the relative control inputs $a_k^{m\star}$ of edge agents.

From how we define the relationship between the relative acceleration control inputs and the real control inputs between the agent in (11) and given the optimal solution (17) to Problem 2, the distributed solution can be obtained by solving

$$\Phi u_k^\star = a_k^\star \tag{21}$$

for $u_k^\star$, where $a_k^\star = [a_k^{1\star^T}, ..., a_k^{M\star^T}]^T$ and $\Phi = \big((-D^T)\otimes I_n\big)$.

Remark 1: We aim to solve (21) in a distributed manner. We replace the problem of minimizing the residual $f(u) = ||\Phi u^\star - a^\star||^2$. Applying the steepest descent gives us an iteration of the form

$$u_{t+1}^\star = (I - 2\alpha\Phi^T\Phi)u_t^\star + 2\alpha\Phi^T a_k^\star, \tag{22}$$

where α is a small learning rate, and t indicates the running index of iterations. It can be shown that this converges to a solution $u_k^\star$ if the learning rate α is small enough and if the solution exists.

Thus the solution (21) can be obtained in a distributed manner, where each agent solves a problem based on locally available information only.

C. A Receding Horizon Manner

The distributed solution to the multi-agent consensus problem presented above can be implemented in a receding horizon manner, i.e. at time k each agent calculates the optimal control sequence over the prediction horizon, but only the first control input is applied and at time $k+1$ the problem is solved again, based on the information available at that time. In fact the optimal gain can be precomputed; the optimal control action on edge i at time k is

Algorithm 1 Optimal control via distributed framework

Input: relative state z at current time k
Output: relative control input a_k, relative states z_{k+1}, transformed control input u_k, and transformed states x_{k+1}

Initialization $\tilde{P}_T = \tilde{Q}_T$ and number of iteration $\bar{I}$
for $k = 1 : T-1$ **do**
  for $j = T-1 : -1 : 1$ **do**
    $\tilde{P}_j^m = \tilde{Q}^m + \tilde{F}^T\tilde{P}_{j+1}^m\tilde{F} + \tilde{F}^T\tilde{P}_{j+1}^m\tilde{G}^m\tilde{K}_j^m$
  end for
  $\tilde{K}_k^m = -(\tilde{R}^m + \tilde{G}^m\tilde{P}_{k+1}^m\tilde{G}^m)^{-1}\tilde{G}^{m^T}\tilde{P}_{k+1}^m\tilde{F}$
  $a_k^m = \tilde{K}_k^m z_k$
  $z_{k+1} = \tilde{F}z_k + \tilde{G}a_k$
  $t = 1$
  while $t \le \bar{I}$ **do**
    $u_{t+1}^i = u_t^i -$
      $2\alpha\sum_{j=1}^N \Phi(j,(ni-1):ni)\ \Phi((ni-1):ni,j)u_t^i$
        $+2\alpha\sum_{j=1}^N \Phi((ni-1):ni,j)a_k^j$
    $t = t+1$
  end while
  $x_{k+1} = Fx_k + \sum_{i=1}^N G^i u_k^i$
end for

$$a_k^{m\star} = \tilde{K}_0^m z_k, \quad m = 1,...,M, \tag{23}$$

where

$$\tilde{K}_0^m = -\Big(\tilde{R}^{mm} + \tilde{G}^{m^T}\tilde{P}_0^m\tilde{G}^m\Big)^{-1}\tilde{G}^{m^T}\tilde{P}_0^m\tilde{F}, \tag{24}$$

and $\tilde{P}_0^m$ is obtained from solving backward the Riccati equation (19).

The closed-loop edge system is then

$$z_{k+1}^\star = \Big(\tilde{F} + \sum_{m=1}^M \tilde{G}^m\tilde{K}_0^m\Big)z_k^\star = \hat{F}z_k^\star, \qquad \text{for} \quad k = 0,...,T. \tag{25}$$

Theorem 3: Suppose the following assumptions hold for the terminal cost $\tilde{Q}_T$ and the terminal controller $\tilde{K}_T$

1. $(\tilde{Q}_T, \tilde{F})$ is detectable
2. $(\tilde{F} + \tilde{G}\tilde{K}_T)^T\tilde{Q}_T(\tilde{F} + \tilde{G}\tilde{K}_T) - \tilde{Q}_T \le -\tilde{Q} - \tilde{K}_T^T\tilde{R}\tilde{K}_T$

where $\tilde{Q}_T = \sum_{m=1}^M \tilde{Q}_T^m$, $\tilde{Q} = \sum_{m=1}^M \tilde{Q}^m$, $\tilde{K}_T = [\tilde{K}_T^1, ..., \tilde{K}_T^M]^T$, $\tilde{R} = \text{blkdiag}(\tilde{R}^{mm})$ with $\tilde{G} = [\tilde{G}^1, ..., \tilde{G}^M]$. The receding horizon controller $\tilde{K}_0^m$ (24) that satisfies the Riccati solution (19) will guarantee asymptotic stability.

Proof: We define $\tilde{K} = [\tilde{K}^1, ..., \tilde{K}^M]^T$ and rearrange the cost (16) as

$$\tilde{J}_k = \frac{1}{2}\Big(Z_k^T\hat{Q}Z_k + A_k^T\hat{R}A_k\Big), \tag{26}$$

where $\hat{Q} = \sum_{m=1}^M \hat{Q}^m$, $\hat{R} = \sum_{m=1}^M \hat{R}^{mm}$, and $A_k = [a_k, a_{k+1}, ..., a_{k+T-1}]^T$. We show that the optimal cost (26) $\tilde{J}_k^\star$ is strictly decreasing, i.e. $\tilde{J}_{k+1}^\star < \tilde{J}_k^\star$. At time k, the optimal relative control sequence is $A_k^\star =$

$[a_k^\star, a_{k+1}^\star, ..., a_{k+T-1}^\star]^T$ and the relative state sequence $Z_k^\star = [z_{k+1}^\star, z_{k+2}^\star, ..., z_{k+T}^\star]^T$. Assume at time $k+1$ we choose the sub-optimal relative control sequence $A_{k+1}^\star = [a_{k+1}^\star, ..., a_{k+T-1}^\star, \tilde{K}_T z_{k+T}^\star]^T$, yielding the state sequence $Z_{k+1} = [z_{k+2}^\star, ..., z_{k+T}^\star, (\tilde{F} + \tilde{G}\tilde{K}_T) z_{k+T}^\star]^T$. Using the notation $||z||_Q^2 = z^T Q z$, we have

$$\begin{aligned} \tilde{J}_{k+1} = &||z_{k+2}^\star||_{\tilde{Q}}^2 + ... + ||z_{k+T}^\star||_{\tilde{Q}}^2 \\ &+ ||(\tilde{F} + \tilde{G}\tilde{K}_T) z_{k+T}^\star||_{\tilde{Q}_T}^2 + ||a_{k+1}^\star||_{\tilde{R}}^2 + ... \\ &+ ||a_{k+T-1}^\star||_{\tilde{R}}^2 + ||\tilde{K}_T z_{k+T}^\star||_{\tilde{R}}^2 \end{aligned} \tag{27}$$

$$\begin{aligned} \tilde{J}_k^\star = &||z_{k+1}^\star||_{\tilde{Q}}^2 + ||z_{k+2}^\star||_{\tilde{Q}}^2 + ... + ||z_{k+T}^\star||_{\tilde{Q}_T}^2 \\ &+ ||a_k^\star||_{\tilde{R}}^2 + ||a_{k+1}^\star||_{\tilde{R}}^2 + ... + ||a_{k+T-1}^\star||_{\tilde{R}}^2 \end{aligned} \tag{28}$$

Comparing (27) and (28) we have

$$\begin{aligned} \tilde{J}_k^\star = &\tilde{J}_{k+1} + ||z_{k+1}^\star||_{\tilde{Q}}^2 + ||a_k^\star||_{\tilde{R}}^2 \\ &- ||(\tilde{F} + \tilde{G}\tilde{K}_T) z_{k+T}^\star||_{\tilde{Q}_T}^2 + ||z_{k+T}^\star||_{\tilde{Q}_T}^2 \\ &- ||z_{k+T}^\star||_{\tilde{Q}}^2 - ||\tilde{K}_T z_{k+T}^\star||_{\tilde{R}}^2 \end{aligned} \tag{29}$$

From the second assumption from the theorem, the last two rows in (29) are non-negative, and thus we have

$$\tilde{J}_{k+1} \leq \tilde{J}_k^\star - ||z_{k+1}^\star||_{\tilde{Q}}^2 - ||a_k^\star||_{\tilde{R}}^2$$

Since $\tilde{J}_{k+1}$ is sub-optimal, we know that $\tilde{J}_{k+1}^\star \leq \tilde{J}_{k+1}$. In other words, we have

$$\tilde{J}_{k+1}^\star < \tilde{J}_{k+1} + ||z_{k+1}^\star||_{\tilde{Q}}^2 + ||a_k^\star||_{\tilde{R}}^2 \leq \tilde{J}_k^\star \tag{30}$$

This proves that $\tilde{J}_k^\star$ is a decaying sequence, hence z_k will converge to the origin. ■

IV. ILLUSTRATIVE EXAMPLE

In this section we illustrate the proposed approach with a consensus control problem where double integrator agents are moving in 2D space. We consider four agents with undirected communication topology, as shown in Fig. 1. We have $N = 4$ and $M = 4$, the incidence matrix is

$$D = \begin{bmatrix} -1 & 0 & 0 & -1 \\ 1 & -1 & 0 & 0 \\ 0 & 1 & -1 & 1 \\ 0 & 0 & 1 & 0 \end{bmatrix}. \tag{31}$$

We assume that all agents have the same initial velocities but different initial positions

$$p_0^1 = \begin{bmatrix} 0.5 \\ 5 \end{bmatrix}, p_0^2 = \begin{bmatrix} 0 \\ 0.5 \end{bmatrix}, p_0^3 = \begin{bmatrix} 5 \\ 4 \end{bmatrix}, p_0^4 = \begin{bmatrix} 4 \\ 0 \end{bmatrix}, \quad v_0^i = \begin{bmatrix} 0 \\ 1 \end{bmatrix}.$$

From the incidence matrix D in (31), we have the initial relative dynamics $z_0 = (I_2 \otimes (-D^T) \otimes I_n) x_0$ as

$$\begin{aligned} z_0^1 &= x_0^1 - x_0^2, & z_0^3 &= x_0^3 - x_0^4, \\ z_0^2 &= x_0^2 - x_0^3, & z_0^4 &= x_0^1 - x_0^3. \end{aligned}$$

We define the weighting matrices for the cost in (16) as $\tilde{W}^m = I_M$. The matrix $\tilde{Q}^m$ is then formed with terminal weighting cost $\tilde{W}_T^m$ with given $\tilde{\eta}^m = 1$ for all agent m. The control weighting matrix is chosen as $\tilde{R}^{ii} = \delta I_{Mn}$, and

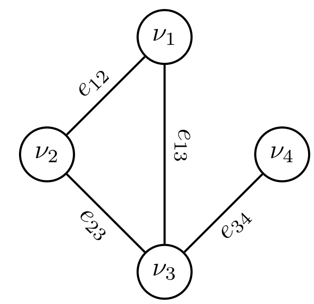

Fig. 1. Undirected graph with $N = 4$ and $M = 6$

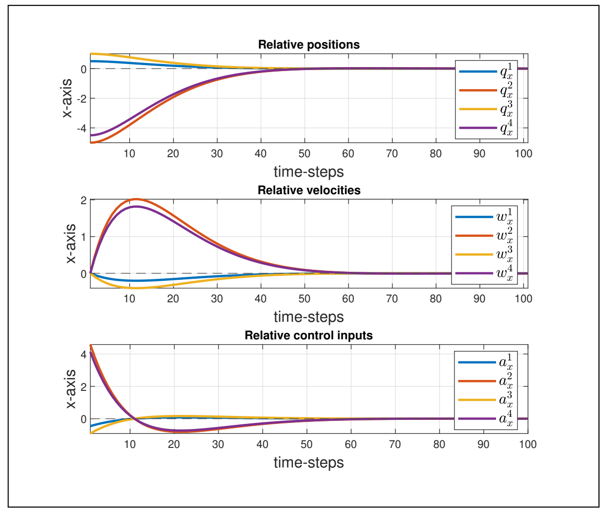

Fig. 2. Response of relative positions q_k^m, relative velocities w_k^m, and relative control inputs $a_k^{m\star}$ in x-direction from Theorem 2

a step size is $\delta = 0.1$. The edge optimization problem can finally be done with taking $T = 10$.

Fig. 2 shows the relative positions, relative velocities, and relative control inputs in x-plane. As we can see here, the relative positions, velocities, and relative control inputs of each agent in the x-direction (as well as in the y-direction) tend to move to 0 when $k \to \infty$. From the given $a_k^{m\star}$, we then calculate $u_k^{i\star}$ iteratively to obtain the distributed solution. We used the learning rate $\alpha = 0.1$ and did iterations 2 and 50 times. To run the iterative scheme, we used the warm-start idea where we initialized $u_t^i = 0$ for $t = 1$. Then for the next iteration, we occupy the last value after being iterated as an initialized value.

To illustrate the convergence speed of the iterative scheme, we run the Algorithm 1 and compare the results with the centralized solution from $u^\dagger = \Phi^\dagger a_k^\star$, where $\Phi^\dagger$ is the pseudo-inverse of Φ. In Fig. 3, the blue stars and red diamonds indicate the solution of the distributed approach and the centralized solution occupying the pseudo-inverse map, respectively. With two iterations only, the control input of Agent 1 in the x-plane converges fast to the centralized solution.

Fig. 4 displays the convergence behavior of the control input of Agent 1 in the x-direction from the first iteration until 50 iterations. As we can see in the figure, the solution of

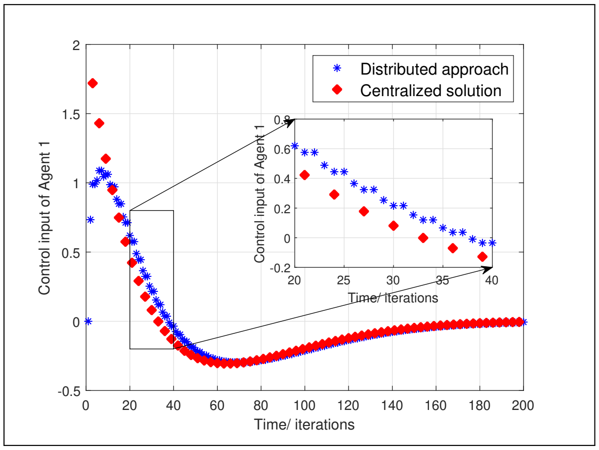

Fig. 3. Convergence of the distributed solution in 2 iterations of Agent 1

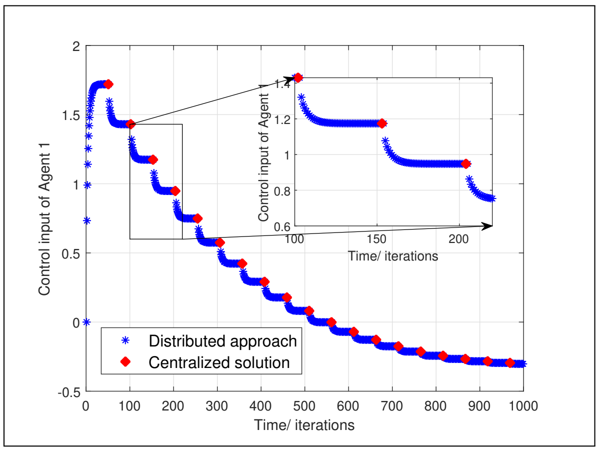

Fig. 4. Convergence of the distributed solution in 50 iterations of Agent 1

the iterative scheme matches the centralized solution (given by the red diamond obtained by solving the pseudo-inverse $\Phi^{\dagger}$) even before 50 iterations are executed.

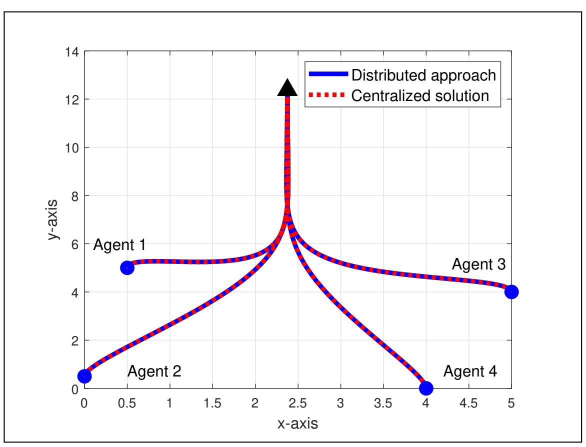

Fig. 5. Trajectories resulting from the distributed approach (blue) and from the centralized solution (red) plotted in $x-y$ plane

The resulting trajectories in the $x-y$ plane under the centralized solution and from the distributed framework are shown in Fig. 5 by the blue and red lines, respectively. As we can see, each agent can reach a consensus as assigned from Problem 1 after the number of time steps evaluated. The progression obtained by the centralized and distributed solution is matched, as can be seen in the figure.

The code for simulation can be found in [10].

V. CONCLUSIONS

In this paper we propose a distributed solution to the LQDTG formulation of a multi-agent consensus problem, which is obtained by solving an auxiliary problem evolving on the edges of the underlying communication graph. In contrast to the original problem, the auxiliary problem can be solved in a distributed manner by each agent using only locally available information. When the terminal cost weights are suitably chosen, the solution obtained from the auxiliary problem is shown to be identical to that of the original problem.

REFERENCES

[1] G. Coulouris, J. Dollimore, and T. Kindberg, *Distributed Systems (3rd Ed.): Concepts and Design*. USA: Addison-Wesley Longman Publishing Co., Inc., 2000.

[2] P. Frihauf, M. Krstic, and T. Basar, “Nash equilibrium seeking in noncooperative games,” *IEEE Transactions on Automatic Control*, vol. 57, no. 5, pp. 1192–1207, 2012.

[3] J. Engwerda, *LQ Dynamic Optimization and Differential Games*. John Wiley & Sons, 2005.

[4] G. Jank and H. Abou-Kandil, “Existence and uniqueness of open-loop nash equilibria in linear-quadratic discrete time games,” *IEEE Transactions on Automatic Control*, vol. 48, no. 2, pp. 267–271, 2003.

[5] D. Gu, “A differential game approach to formation control,” *IEEE Transactions on Control Systems Technology*, vol. 16, no. 1, pp. 85–93, 2008.

[6] J. H. Barghi and N. Vasif, “On the finite horizon nash equilibrium solution in the differential game approach to formation control,” *Journal of Systems Engineering and Electronics*, vol. 30, no. 6, pp. 1233–1242, 2019.

[7] H. B. Jond and J. Platoš, *Autonomous Vehicle Convoy Control as a Differential Game*, 2021. [Online]. Available: arXiv:2101.08858v2

[8] F. M. Callier and C. A. Desoer, *Linear system theory / Frank M. Callier, Charles A. Desoer*. Springer-Verlag New York, 1991.

[9] F. Bullo, *Lectures on Network Systems*, 1st ed. Kindle Direct Publishing, 2021. [Online]. Available: http://motion.me.ucsb.edu/book-lns

[10] P. Aditya and H. Werner, *Code for the paper "A Distributed Linear Quadratic Discrete-Time Game Approach to Multi-Agent Consensus*. Zenodo, 2022. [Online]. Available: https://doi.org/10.5281/zenodo.6402300