2015 International Conference on Unmanned Aircraft Systems (ICUAS)
Denver Marriott Tech Center
Denver, Colorado, USA, June 9-12, 2015

Multiple Cooperative UAVs Target Tracking using Learning Based Model Predictive Control

Ahmed. T. Hafez[1] , Sidney. N. Givigi [2], Khaled A. Ghamry[3] and Shahram. Yousefi[1]

***Abstract*—In this paper, formation of a group of multiple cooperative unmanned aerial vehicles (UAVs) in a desired geometrical pattern while tracking an aerial target is implemented using decentralized Learning Based Model Predictive Control (LBMPC). The LBMPC is a new control technique that combines statistical learning along with control engineering providing guarantees on safety, robustness and convergence. The controller derived in this paper demonstrates the ability of the vehicles to cooperate, in a decentralized manner, to solve the formation problem in the presence of system uncertainties. The proposed controller respects the general formation constraints known as Reynolds rules of flocking during simulations. Our main contribution in this paper lays in the use of decentralized LBMPC in solving the problem of formation for a group of cooperative UAVs tracking an aerial target in the presence of unmodeled dynamics. A theoretical proof for stability will support our proposed controller.**

***Index Terms*—Learning Based Model Predictive Control, Unmanned Aerial Vehicles,Convex Optimization, Cooperative Robotics**

I. Introduction

Multiple UAVs cooperate using general strategies known as UAV tactics. UAV tactics are defined as the procedure used by the UAV team to execute their required missions. These strategies can be either centralized, as in the case where a group of UAVs receives coordinated instructions from one centralized decision maker; or decentralized, in which the UAVs are responsible for making their individual decisions. UAV tactics can be classified into: swarming or formation, task assignment, formation reconfiguration and dynamic encirclement [1].

Formation is defined as the ability of a group of cooperative vehicles to choose their own location with a pre-determined geometrical patterns [1]. UAVs using formation tactic are governed by Reynolds Rules of Flocking which present the behavior required for birds [2]:

- Collision Avoidance: It means that each UAV in the team has to avoid collisions with the nearby flock mates
- Velocity Matching: It means that each UAV in the team attempts to match its velocity with the nearby flock mates
- Flock Centering: It means that each UAV in the team attempts to stay close to its nearby flock mates

In order to avoid failure of the mission, formation for cooperative UAVs has different approaches according to external factors affecting the UAV team. These factors can be summarized as [3]:

- Changing the position of UAVs during flight.
- Combining small groups to form a large group according to the required mission.
- Breaking a large group into smaller groups to perform more than one application at the same time.

Some of the main formation structure approaches are leader follower approach [4], virtual structure approach [5] and behavior- based approach [6]. There is an extensive amount of research in the field of formation for multiple cooperative UAVs. In the presence of bad weather conditions, obstacles, aerial jamming on the communication channels between UAV team members and external threats which prohibit the team from performing their required mission, formation tactic appears as an optimal solution for the UAV team to fulfill their mission by switching from a particular geometric pattern to another [7, 8, 9, 10].

For instance, a formation of three UAVs is controlled by a centralized multi-layer control scheme to follow a predetermined trajectory in a tracking missions. Each layer in the control scheme is responsible to generate the required control input for each vehicle [9]. In [10], a robust control algorithm accompanied with a higher level path generation method is used to control the structure of a group of cooperative UAVs. The goal is to perform formation change maneuvers with a guaranteed safe distance between the different members of the team throughout the whole mission. The robust control ensures the stability of the formation during maneuver while the path generation method provides the vehicle with the safe paths.

The design of a controller to modify the response and behavior of a system to meet certain performance requirements is a challenging problem in many control applications. Model Predictive Control (MPC) is characterized by its ability to handle difficult constraints on control inputs and states, and to deal with multi-variable, multi-input multi-output systems [11].

For UAV formation, MPC is an optimum solution to solve the problem for a group of cooperative autonomous vehicles in different formation reconfiguration [12, 13, 14].

[1]Ahmed. T. Hafez and Shahram. Yousefi are with the Department of Electrical and Computer Engineering, Queens University, Kingston, ON. Canada 11edhk@queensu.ca, s.yousefi@queensu.ca

[2] Sidney. N. Givigi is with the Department of Electrical and Computer Engineering, Royal Military College of Canada, Kingston, ON. Canada sidney.givigi@rmc.ca

[3] Khaled. Ghamry is with the Department of Mechanical and Industrial Engineering,Concordia University, Montréal, QC H3G1M8, Canada. k_ghamry@encs.concordia.ca

978-1-4799-6009-5/15/$31.00 ©2015 Crown

For instance, Nonlinear Model Predictive Control (NMPC) was used to control the formation of a fleet of UAVs in the presence of obstacles and collision avoidance [14], while in [15, 16], an NMPC is used to develop the formation guidance for a team of UAVs, where the controllers predict the behavior of the system over the receding horizon and generate the optimal control commands over the horizon for flight formation and inter-collision avoidance in the presence of control input and state constraints.

Recently, control designers started to investigate the effect of adding a learning algorithm to MPC. Applying a learning algorithm combined with MPC will improve the performance of the system and guarantees safety, robustness and convergence in the presence of states and control inputs constraints. In [17], a quadrotor used LBMPC to learn its non-linear dynamics and stabilize in a desired altitude. A dual extended kalman filter (DEKF) was used in the learning phase, while an MPC was used to solve the optimization control problem. Moreover, LBMPC was used by a single quadrotor in [18] to catch a ball using the same concept in [17].

In [19], an LBMPC, based on the tube MPC technique, was used successfully in controlling a team of three UAVs stabilizing in desired formation while tracking a reference trajectory even when the trajectory changes its directions suddenly. The learning algorithm, a DEKF, succeeded to learn the unmodeled dynamics and generated the correct learning parameters. Moreover, a theoretical proof for the stability of the formation was introduced.

Our main contribution in this paper is to solve the formation problem for a group of cooperative UAVs tracking an aerial target using decentralized LBMPC in the presence of system uncertainties. On one hand, the learning algorithm will be used to learn the uncertainties and the unmodeled dynamics of the system, while on the other hand, the decentralized MPC will minimize the optimization control cost.

The paper is organized as follows. We start with a discussion of the notations used throughout the paper in Section II, followed by the formulation of the problem and the specific control objectives in Section III. In Section IV, we introduce a theoretical analysis of the problem accompanied with proof of stability of the system, while we develop our LBMPC and show its constraints in Section V. In Section VI, we present the results of our simulations. We conclude our work in Section VII and discuss some future works.

II. Preliminaries

In this section, we define the notations used in this paper. Vectors are not typeset specially, but will be identified as such when introduced (e.g., $v \in \mathbb{R}^{10}$). All vectors are column vectors. We use superscript A^T to denote the transpose of a matrix A and v^T to denote the transpose of a vector v.

Variables that change at each discrete time step have the time index denoted by the subscript. However, as we are using a group of cooperative UAVs , the number of the UAV in the team is subscripted (e.g, $x_{i,\tau}$ means the state x of the i^{th} vehicle at time step τ). In equations describing the update of such a variable in the next time step $(\tau + 1)$, the variable is denoted by $x_{i,\tau+1}$.

The notation $\|v\|^2_M$ denotes the quadratic form $v^T M v$. Symbols with a dot above them are the time derivative of that quantity (e.g., $\dot{x} = \frac{d}{dt}x$). Marks above the variable indicate the different models of the same system. For instance, the true system has the state x, the linear nominal system has the state $\bar{x}$, the system with the oracle has the state $\tilde{x}$ and the estimated system has the state $\hat{x}$. Similar marks are used for the corresponding control inputs and outputs.

The Minkowski sum of two sets $\mathcal{U}$ and $\mathcal{V}$ is defined as $\mathcal{U} \oplus \mathcal{V} = \{u + v | u \in \mathcal{U}; v \in \mathcal{V}\}$. Their Pontryagin set difference is defined as $\mathcal{U} \ominus \mathcal{V} = \{u \in \mathcal{U} | u + v \in \mathcal{U}, \forall v \in \mathcal{V}\}$. This set difference is not symmetric, and so the order in which these operations are performed can make a difference. For the set $\mathcal{U}$, the linear transformation of the set by matrix T is given by $T\mathcal{U} = \{Tu : u \in \mathcal{U}\}$ [20, 21, 22].

III. System description and control objectives

We aim in this paper to solve the problem of formation for a group of N cooperative UAVs while tracking a target. The main objective is to stabilize the team toward an equilibrium point in the center of the formation using decentralized LBMPC. In order to simplify the presentation, let us consider that the UAVs and the target act in a two dimensional space such that the height and yaw controllers have no influence on the lateral movement of the vehicles. However, notice that the problem could be extended to the three dimensional case with increased computational demands [23].

A. System Modeling

For each UAV $i \in \{1, ..., N\}$, the states and control inputs are denoted by

$$z_i(t) = \begin{bmatrix} x_i(t) \\ y_i(t) \\ \theta_i(t) \end{bmatrix} \in \mathcal{Z} \subset \mathbb{R}^n \tag{1a}$$

$$u_i(t) = \begin{bmatrix} V_i(t) \\ \omega_i(t) \end{bmatrix} \in \mathcal{U} \subset \mathbb{R}^m \tag{1b}$$

where x_i and y_i is the position of the i^{th} UAV in 2D- Cartesian frame and θ_i is the heading angle. Considering that the UAVs are moving in the horizontal plane, the dynamics of the i^{th} UAV is described by:

$$\dot{z}_i(t) = f(z_i(t), u_i(t)) = \begin{bmatrix} V_i(t) cos\theta_i(t) \\ V_i(t) sin\theta_i(t) \\ \omega_i(t) \end{bmatrix} \tag{2}$$

where V_i is the linear velocity and ω_i is the angular turning rate of the i^{th} UAV. The control inputs of the i^{th} UAV is denoted by $u_i \in \mathcal{U}$, while $z_i \in \mathcal{Z}$ is the measurable states. We assume that the target has the same dynamic model as the vehicles but denoted by k instead of i.

The dynamic model state space representation of each vehicle is given as:

$$\dot{z}_i(t) = A_i z_i(t) + B_i u_i(t) + g(z_i, u_i) \tag{3a}$$

$$Y_i(t) = C_i z_i(t) + \epsilon \tag{3b}$$

where, again, z_i is the states of the i^{th} UAV in the team, u_i is its control inputs, Y_i is the measured outputs of the i^{th} UAV, $g(z_i, u_i)$ is the unmodeled dynamics of the i^{th} UAV and ϵ represents the measurement noise, assumed to be a bounded stochastic quantity. We assume that the modeling error $g(z_i, u_i)$ is bounded and lies within a polytope $\mathcal{W}$ such that $g(z_i, u_i) \in \mathcal{W}\ \forall (z_i, u_i) \in (\mathcal{Z}, \mathcal{U})$. Thus the nominal dynamic state (the case in which $g(z_i, u_i) \equiv 0$) is,

$$\dot{\bar{z}}_i(t) = \mathcal{F}_\tau(z_i, u_i) = A_i \bar{z}_i(t) + B_i \bar{u}_i(t) \tag{4}$$

B. Problem Definition

In this section, Learning Based Model Predictive Control (LBMPC) is used to solve the problem of formation of a group of cooperative UAVs in the presence of state and control inputs constraints. LBMPC seeks to combine attributes of MPC (most notably, the ability to enforce constraints, which encode safety requirements) with elements of adaptive or learning schemes which promise to improve performance by improving system models based on data obtained on-line.

For each vehicle $i \in \{1, ..., N\}$, the unmodeled dynamics $g(z_i, u_i)$ for the i^{th} UAV is represented by a linear, time-varying oracle $\mathcal{O}_\tau : \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}^n$, parametrized by a vector of parameters β. The state space form of the oracle is given as:

$$\mathcal{O}_\tau(\tilde{z}_i, \tilde{u}_i) = F_i(\beta)\tilde{z}_{i,\tau} + H_i(\beta)\tilde{u}_{i,\tau} \tag{5}$$

where $\mathcal{O}_\tau(\tilde{z}_i, \tilde{u}_i) \in \mathcal{W}\ \ \forall\ \ (\tilde{z}_i, \tilde{u}_i) \in (\mathcal{Z}, \mathcal{U})$ returns an estimate of the unmodeled dynamics, $F_i(\beta)$ is the oracle updates to the dynamics matrix for the i^{th} vehicle, $H_i(\beta)$ is the oracle updates to control input matrix for the i^{th} vehicle, and β is the true vector learning parameters in the oracle model.

By adding the time varying oracle to the system, we denote the states for the system with an oracle for the i^{th} vehicle in the team as $\tilde{z}, \tilde{y}, \tilde{u}$, and its dynamics are represented as:

$$\begin{aligned} \tilde{z}_{i,\tau+1} &= \mathcal{F}_\tau(z_i, u_i) + \mathcal{O}_\tau(\tilde{z}_i, \tilde{u}_i) \\ &= (A_i + F_i(\beta))\tilde{z}_{i,\tau} + (B_i + H_i(\beta))\tilde{u}_{i,\tau} \end{aligned} \tag{6}$$

Each UAV in the team senses the presence of the traget and estimates the position for the other neighbour vehicles in the team. Each vehicle generates its own control signals, in a decentralized manner, to minimze the optimization cost function and ensure the stability of the group while tracking the target. To show the difference between all the above equations for the i^{th} UAV at time step τ, we need to define the following: $z_{i,\tau}$ is the true system state dynamics , $\tilde{z}_{i,\tau}$ is the state dynamics of the system with the oracle state, while $\hat{z}_{i,\tau}$ is the estimated state dynamics of the system and $\bar{z}_{i,\tau}$ is the state dynamics of the nominal system with disturbance state. The states and control inputs constraints are used to guarantee both feasibility and convergence of the designed control policy [18].

IV. Theoretical Analysis

In this section, we will prove that by applying our proposed controller, the closed loop system (6) converges to the desired equilibrium states and the UAV team will stabilize in the desired formation while tracking the target.

In [24], the problem of formation stabilization was solved using distributed receding horizon control in the presence of a group of constraints on every vehicle i as follows:

- Each vehicle i can estimates the current states of the nieghbours z_i.
- Each vehicle i computes its own optimal control trajectory, comparing it to an assumed control trajectory.
- Before the next update, each vehicle i implements the current optimal control trajectory, computes the next assumed control trajectory, transmits it to all of its neighbors and receives the assumed control trajectories from all the neighbors.

We are interested in this paper in solving the formation problem using LBMPC in the presence of state dynamics constraints and control inputs constraints mentioned in [24] in the presence of unmodeled system dynamics affecting the system every time step.

Our decentralized LBMPC for each vehicle will have a dual extended Kalman filter to learn about its uncertainties, its optimal control inputs, its feedback gain and updating its matrices each time step, an MPC to solve the optimization control problem. The learning phase beside the prediction phase will allow the system to converge to the desired equilibrium point in the presence of the unmodeled dynamics $g(z_i, u_i)$ leading to the stabilization of the UAV formation while tracking the target.

We will begin by mentioning some basic assumptions that support our theoretical analysis.

Assumption 1. The optimal cost function of the LBMPC, in addition to the system states and control inputs vectors are calculated at each time step, and the computational time is in trade-off with the solution of the optimization problem.

Assumption 2. The UAV system is controllable and observable, meaning any state $\bar{z}_i(t) \in \mathcal{Z}$ can be controlled to any other state $\hat{z}_i(t) \in \mathcal{Z}$ in finite time for $i \in \{1, ...N\}$.

Assumption 3. For every vehicle $i \in \{1, ..., N\}$, the unmodeled dynamics $g(z_i, u_i)$ for the i^{th} UAV are represented by a time-vary oracle $\mathcal{O}_\tau(\tilde{z}_i, \tilde{u}_i) = F_i(\beta)\tilde{z}_{i,\tau} + H_i(\beta)\tilde{u}_{i,\tau}$ such that it is a bounded polytopes and lies within $\mathcal{W}$ such that $\mathcal{O}_\tau(\tilde{z}_{i,\tau}, \tilde{u}_{i,\tau}) \in \mathcal{W}\ \forall (\tilde{z}_i, \tilde{u}_i) \in (\mathcal{Z}, \mathcal{U})$.

where $F_i(\beta)$ are the oracle updates to the dynamics matrix for the i^{th} vehicle, $H_i(\beta)$ is the oracle updates to control input matrix for the i^{th} vehicle, and β is the true vector learning parameters in the oracle model.

From the Assumptions (1-3) and by following the same technique as used in [24] in order to show that the multiple vehicle system converges to a stable formation.

Theorem 1. Assume that the problem as defined in section III is feasible at the first time step. Furthermore, the constraints as defined in [24] are satisfied and propagated. Then, if the time step is chosen to be smaller than δ_{max} as defined in Theorem 1 of [24], for any discrete period $k \in (\tau, \tau + M]$, the optimization control problem will have a feasible solution at any update time $j \in \{1, 2, ..., M\}$.

Proof. Since Theorem 1 in [19] supported by Theorem 1 in [21] states that for each one of the distributed problems $i \in \{1, 2, \cdots, N\}$ a sequence $\mathcal{M}_i^0$ may be found such that the problem is feasible, i.e., it may be solved given a sufficiently large horizon. Let us assume now a collection of MPC tubes represented by $\mathcal{R}_1 \times \mathcal{R}_2 \times \cdots \times \mathcal{R}_N$ and each region $\mathcal{R}_i$ includes the stochastic bounds for the unmodeled dynamics $\mathcal{W}$. Having one vehicle depend on others is equivalent to expanding the stochastic bounds represented by the polytope $\mathcal{W}_i$ to include the noise coming from the other vehicles $k \in \{1, 2, ..., N\}$ such that $k \neq i$ and we may represent these new bounds as $\mathcal{W}_i'$. Therefore, we may write the evolution of the regions $\mathcal{R}_i^j$ as

$$\mathcal{R}_i^j = \oplus_{l=0}^{j-1} \ (A_i + B_i K_i)^l \ \mathcal{W}_i'$$

and $\mathcal{R}_i^0 = \{0\}$. Notice again that we are assuming that the dynamics of the system are linear. Therefore, since all the conditions in [21] are satisfied, the individual LBMPC problems are robustly asymptotically stable (RAS), even for the new extended polytopes.

Therefore, if the conditions of Theorem 1 of [24] are satisfied, it will follow that the vehicles will generate trajectories that will stabilize the formation. We assume that the dynamics of the system in (1) are twice continuously differentiable. Also, the local control law for each vehicle $\bar{u}_{i,\tau} = K_i(\bar{z}_{i,\tau} - z_{ref}) + \bar{u}_{ref}$ is Schur stable such that z_{ref} and u_{ref} are the reference states and control inputs, respectively. Theorem 8 in [25] guarantees that the trajectories will be bounded and therefore, Assumption 1 of [24] is satisfied. Finally, our proposed control policy admits a time step as small as necessary. Therefore, choosing our time step $\tau < \delta_{max}$ guarantees that the formation may be achieved by the application of Theorem 1 of [24]. $\square$

V. Control Design

In this section, we will outline the design of a learning based model predictive controller used to solve the formation problem for a group of cooperative UAVs while tracking a target.

The overall control architecture is composed of two main parts:

- Estimation of the vehicles states and learning the unmodeled dynamics.
- Solving the Quadratic Programming (QP) optimization control problem for the closed loop system.

On one hand, the estimation of the system states depends on using a nominal model to make the predictions of the current states based on the past states and control inputs. On the other hand, the optimization problem uses a system model to determine the cost of the control policies and evaluate the result of those policies over a finite planning horizon [18].

A. Learning Approach

Several approaches for learning the system dynamics can be used for the approximation of the uncertain parameters of the system in our proposed controller such as reinforcement learning and various types of Kalman Filters like Extended Kalman Filter (EKF) and Dual Extended Kalman Filter (DEKF). In our paper, the DEKF was chosen because we know the structure of the system (the general form of the equations of motion) and which parameters are uncertain.

The Dual Extended Kalman Filter (DEKF), introduced by Wan and Nelson [26], is an extension that has been developed for the state and parameter estimation of non-linear systems. The main concept in the DEKF is the combination between the state estimation and the parameter estimation using two EKFs in parallel, which allows to switch off the parameter estimator once we reach a sufficiently good set of estimates for the parameters. This should improve the performance of the system by reducing the parameter uncertainties [27].

In our system (3), the unmodeled dynamics $g(z_i, u_i)$ for the i^{th} UAV in the fleet is calculated using the oracle $\mathcal{O}_\tau(\tilde{z}_i, \tilde{u}_i)$ in (5). The parameters of the oracle correspond to the linearization of the unmodeled dynamics about an operating point that varies during the UAV flight. The basic equations used for building the DEKF for each i^{th} vehicle are stated as in [27]

Parameter prediction:

$$\hat{z}_{ip}^-(t) = \hat{z}_{ip}(t-1) \tag{7a}$$

$$P_{ip}^-(t) = P_{ip}^-(t-1) + R_p \tag{7b}$$

State prediction:

$$\hat{z}_i^-(t) = f(\hat{z}_i^-(t-1), u_i(t), \hat{z}_{ip}^-(t)) \tag{8a}$$

$$P_{is}^-(t) = J_s(t) P_{is}^-(t-1) J_s^T(t) + R_s \tag{8b}$$

State Correction:

$$K_s(t) = P_{is}^-(t) H_s^T [\sigma_s + H_s P_{is}^-(t) H_s^T]^{-1} \tag{9a}$$

$$\hat{z}_i(t) = \hat{z}_i^-(t) + K_s(t)[y_i(t) - H_s \hat{z}_i^-(t)] \tag{9b}$$

$$P_{is}(t) = [I - K_s(t) H_s] P_{is}^-(t) \tag{9c}$$

Parameter Correction:

$$K_p(t) = P_{ip}^-(t) H_p^T [\sigma_p + H_p P_{ip}^-(t) H_p^T]^{-1} \tag{10a}$$

$$\hat{z}_{ip}(t) = \hat{z}_{ip}^-(t) + K_p(t)[y_i(t) - H_s \hat{z}_i^-(t)] \tag{10b}$$

$$P_{ip}(t) = [I - K_p(t) H_p] P_{ip}^-(t) \tag{10c}$$

where $\hat{z}_i$ is the estimated state vector, $\hat{z}_{ip}$ the estimated parameter vector, u_i is the control input vector and y_i is the measurement vector for the i^{th} vehicle. The error covariance matrix of the state vector for each vehicle is denoted by P_{is}, while the error covariance matrix of the parameter vector is denoted by P_{ip}. Also, R_s and R_p are the user specified process noise covariance matrices for the state and parameter estimators, respectively, while, the corresponding output noise covariance matrices for the state and parameter estimators are denoted by σ_s and σ_p, respectively. Moreover, K_s and K_p are the DEKF gain matrices for the state and the parameter, respectively, while H_s and H_p are the Jacobian matrices of the output for state/parameter estimator and J_s is the Jacobian matrices for state estimates.

B. Optimization Design

The proposed controller uses the idea of tube-MPC to ensure that the true trajectory for each vehicle i is close enough to its nominal trajectory. Also, by the same argument, the predicted trajectories of the neighbors are near enough from their true trajectories. The effect of the disturbance for each vehicle i is controlled using the feedback gain K_i.

For the LBMPC prediction horizon M, given a nominal trajectory, the true trajectory for each agent i is guaranteed to lie within an MPC tube around the given trajectory. The width of the MPC tube at the j^{th} time step, for $j \in l = \{0, ..., M-1\}$, is given by a set $\mathcal{R}_j$, while the width of the MPC tube at M is given by $\mathcal{R}_M$. The state constraints $\mathcal{Z}$ are shrunk by the width of this tube, which means that if the given nominal trajectory lies in $\mathcal{Z} \ominus \mathcal{R}_j$, then the true trajectory lies in $\mathcal{Z}$.

At the heart of the LBMPC control scheme is the on-line solution of a convex optimization problem. The optimization cost function will contain the learning part for each vehicle in the team

$$\min_{(c_i)} J_i(\tilde{z}_{i,[\tau:\tau+M]}, \hat{z}_{-i,[\tau:\tau+M]}, \check{u}_{i,[\tau:\tau+M]}) \tag{11}$$

subject to:

$$\tilde{z}_{i,\tau} = z_{i,\tau}, \qquad \bar{z}_{i,\tau} = z_{i,\tau}$$

$$\tilde{z}_{i,\tau+j} = A_i \tilde{z}_{i,\tau+j-1} + B_i \check{u}_{i,\tau+j-1} + \mathcal{O}_\tau(\tilde{z}_i, \check{u}_i)$$

$$\left\{ \begin{array}{l} \bar{z}_{i,\tau+j} = A_i \bar{z}_{i,\tau+j-1} + B_i \check{u}_{i,\tau+j-1} \\ \check{u}_{i,\tau+j-1} = K_i \bar{z}_{i,\tau+j-1} + c_{i,\tau+j-1} \\ \bar{z}_{i,\tau+j} \in \mathcal{Z} \ominus \mathcal{R}_j, \quad \check{u}_{i,\tau+j-1} \in \mathcal{U} \ominus K_i \mathcal{R}_j \\ \bar{z}_{i,\tau+M} \in \mathcal{Z}_M \ominus \mathcal{R}_M \end{array} \right. \tag{12}$$

where M is the prediction horizon, $\tilde{z}_{i,\tau}$ is the predicted dynamic states of the i^{th} vehicle while $\hat{z}_{-i,\tau}$ denote the vector of the estimated dynamic states of the neighbors of i, $\check{u}_{i,\tau}$ is the optimal control inputs to the system and $c_{i,\tau+j-1}$ is the set of control inputs generated by the MPC. The polyhedral sets $\mathcal{Z}$ and $\mathcal{U}$ are bounded and convex; they encode the allowable states and control inputs, respectively.

The cost functions J_i are non negative functions of the states and use the oracle $\mathcal{O}_\tau(\tilde{z}_i, \check{u}_i)$ to update the states for each i^{th} vehicle and are given by

$$J_i = p(\tilde{z}_{i,\tau+M}, \hat{z}_{-i,\tau+M}) + \sum_{j=0}^{M-1} q(\tilde{z}_{i,\tau+j}, \hat{z}_{-i,\tau+j}) + r(\check{u}_{i,\tau+j}) \tag{13}$$

where $p(\tilde{z}_{i,\tau+M}, \hat{z}_{-i,\tau+M}) = \|\tilde{z}_i - \hat{z}_{-i} - z_s\|_P^2$ is the final state error cost, $q(\tilde{z}_{i,\tau+j}, \hat{z}_{-i,\tau+j}) = \|\tilde{z}_i - \hat{z}_{-i} - z_s\|_Q^2$ is the intermediate step costs and the control input cost is $r(\check{u}_i) = \|\check{u}_i - u_s\|_R^2$. Matrices P, Q, and R are semi-definite positive matrices weights on the final state error cost, the intermediate state error cost and the control input cost, respectively.

Moreover, the desired state is denoted by z_s, while u_s is the steady state control that would maintain the desired steady state z_s, the nominal feedback gain K_i serves to limit the effect of model uncertainty for agent i and it is chosen so that the discrete-time algebraic Ricatti equation (DARE)

$$(A_i + B_i K_i)^T P (A_i + B_i K_i) - P = (Q + K_i^T R K_i)$$

is satisfied. While $\mathcal{R}_0 = \{0\}$, $\mathcal{R}_j = \oplus_{l=0}^{j-1} (A_i + B_i K_i)^l \mathcal{W}$ and $\mathcal{R}_M$ is the width of the MPC tube at the M^{th} step.

Finally, the actual control inputs $\check{u}_{i,\tau}$ in (12) are used to determine the predicted learning dynamic states $\tilde{z}_{i,\tau}$ in (11) and the predicted nominal state $\bar{z}_{i,\tau}$ used for constraint satisfaction in (12) for the i^{th} vehicle.

VI. Simulation Results

The control strategy discussed in sections III, IV and V is successfully implemented in simulation by a multi-UAV team consisting of three vehicles. The objective of these simulations is to show that, the LBMPC policy designed is fit for solving the problem of formation for a group of cooperative UAVs while tracking a target. The state and control input vectors are defined as:
$\mathcal{Z} = \mathbb{R}^3$, $\quad \mathcal{U} = \{(V_i, \omega_i) \in \mathbb{R}^2 : -1 \leq V_i \leq 1, -0.5 \leq \omega_i \leq 0.5\}$. By linearizing (2) using the Jacobian, (3) can be written as:

$$\dot{\bar{z}}_i = \begin{bmatrix} 0 & 0 & -V_i sin\theta_i \\ 0 & 0 & V_i cos\theta_i \\ 0 & 0 & 0 \end{bmatrix} \bar{z}_i + \begin{bmatrix} cos\theta_i & 0 \\ 0 & -sin\theta_i \\ 0 & 1 \end{bmatrix} \bar{u}_i \tag{14}$$

At the first time step, we consider θ_i is equal to 0 rad, then the matrices A_i and B_i are given by:

$$A_i = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix}; \quad B_i = \begin{bmatrix} 1 & 0 \\ 0 & 0 \\ 0 & 1 \end{bmatrix}$$

The aim of our simulations is the success of a group of cooperative UAVs to form a desired formation while tracking the target $z_k = (x_k(t), y_k(t), \theta_k) \in \mathbb{R}^3$. For the compatibility with our objective, we rewrite the system dynamics in (14) in the error form.

For a UAV team consists of 3 vehicles, the error system is given as:

$$\begin{bmatrix} \dot{E}_{12} \\ \dot{E}_{13} \\ \dot{E}_{23} \end{bmatrix} = \begin{bmatrix} A & 0 & 0 \\ 0 & A & 0 \\ 0 & 0 & A \end{bmatrix} \begin{bmatrix} E_{12} \\ E_{13} \\ E_{23} \end{bmatrix} + \begin{bmatrix} B & -B & 0 \\ B & 0 & -B \\ 0 & B & -B \end{bmatrix} \begin{bmatrix} U_1 \\ U_2 \\ U_3 \end{bmatrix} \tag{15}$$

In our simulations, the three UAVs are located along the Y axis, where UAV 1 is located at the origin, UAV 2 is located at (0,12) and UAV 3 is located at (0,-12) with initial control inputs zero for the all agents. The target is located at (5,5). The prediction horizon $M = 5$, the simulation time $T = 30$ sec, and the sampling time $\tau = 0.2$ sec.

The requirements of our simulations are based on Reynolds rules and can be summarized as follows:

- Separating distance $d = 10m$ between each two neighbor vehicles in X-Y directions.
- Each vehicle attempts to match its velocity with its neighbors.
- Each vehicle in the fleet attempts to avoid collision with the nearby vehicles.
- Separating distance between the target and the center of the UAV formation $d_s = 5m$

A team of three cooperative UAVs will form a line-abreast formation during tracking a target moving along the $X - Y$ Cartesian Coordinate. The target trajectory is defined as:

$$z_k(t) = \begin{bmatrix} x_k(t) \\ y_k(t) \\ \theta_k(t) \end{bmatrix} = \begin{bmatrix} t \\ t \\ 0 \end{bmatrix} \tag{16}$$

In our simulation, the three UAVs successfully form a line-abreast formation in the X direction with a desired distance of 10m between each two neighbors in the Y direction. The three UAVs track the target moving along the $X - Y$ direction. At each time step, each UAV updates its dynamics matrices using the DEKF to learn about its unmodeled dynamics, estimates the trajectories of the neighbours vehicles and generates its control command in a decentralized manner. Despite the fact that collision avoidance is not incorporated in our optimal control problem (11), either by cost or constraint, we have observed no collision occur between any two vehicles during all our simulations.

The path of the three UAVs and the target trajectory is presented in Fig.1, where UAV1 is presented in a red color, UAV2 with blue color, UAV3 with green color, the target trajectory with black and the average position of the three vehicles $\bar{z}_{\sum}(t)$ is denoted with magenta. The optimal control inputs (c_i) for each agent in the fleet are presented in Fig. 2, where all the c_i control inputs generated by the MPC of each vehicle converge to zero as the vehicles are stabilized in their desired formation while tracking the target. The separating distance between UAV1 and its neighbors are presented in Fig. 3, where the separating distances converge to the desired distance $d = 10m$, while the separating distance between the target position and the center of the fleet is shown in Fig. 4, where it converges to the desired separating distance.

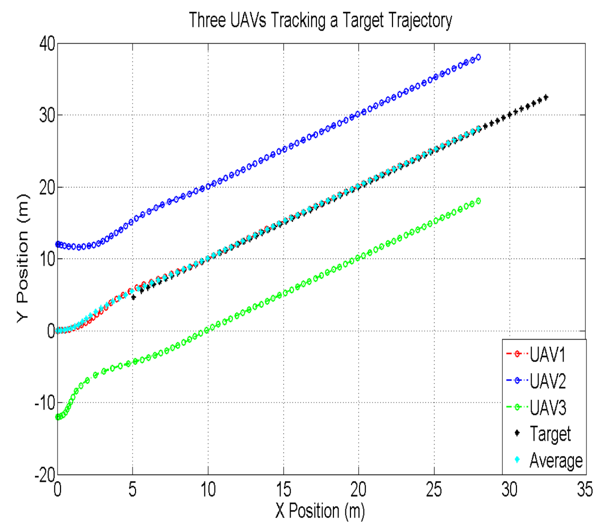

Fig. 1. The three UAVs form a line-abreast formation with a desired separation distance $d = 10$ m while tracking the target trajectory

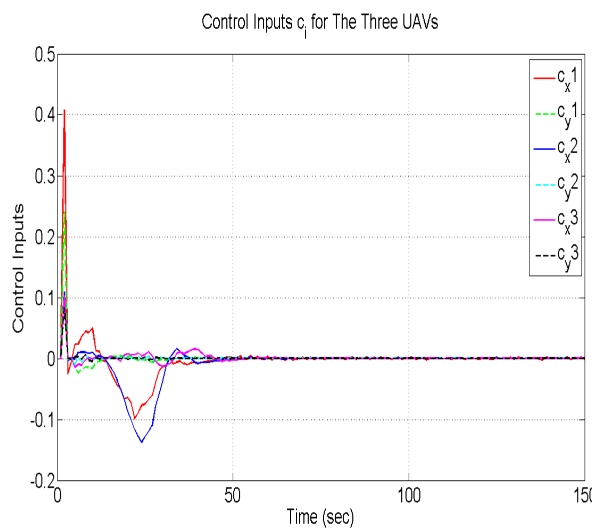

Fig. 2. The optimal control inputs for the three agents, the control inputs c_i converge to zero due to the stabilization of the desired formation while tracking the target trajectory.

VII. Conclusion

In this paper, we have described the design and implementation of a decentralized LBMPC controller for a team of cooperative UAVs to track a reference trajectory with a desired formation. We have shown that our system succeed to learn the unmodeled dynamics using the learning algorithm and converge to the desired results. Also, our results show that applying decentralized LBMPC strategy

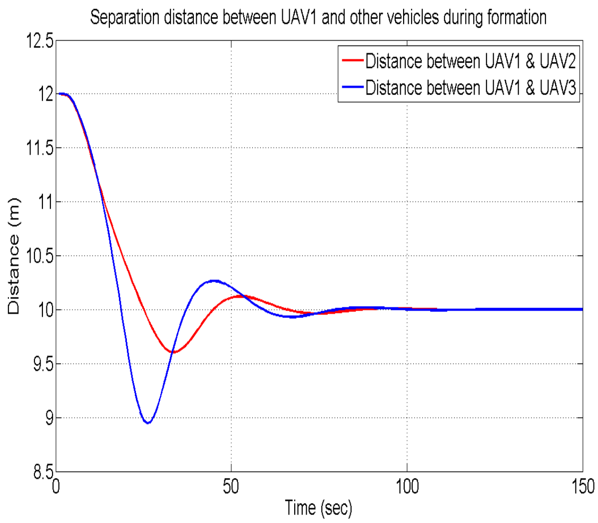

Fig. 3. The separation distance between UAV1 and other vehicles in the fleet.

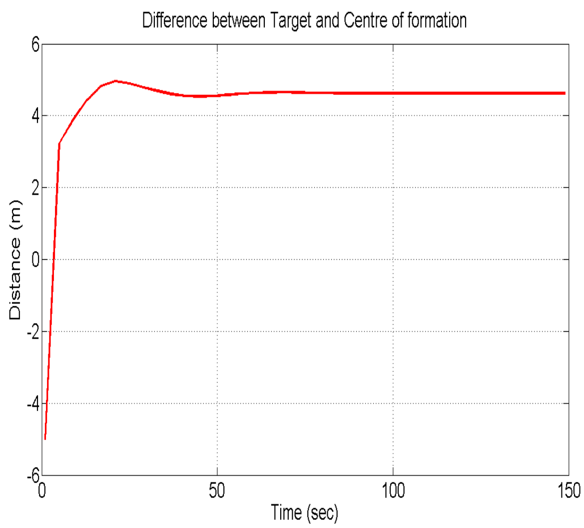

Fig. 4. The separation distance between the target position and the center of the fleet.

succeed to solve the problem of formation for a group of cooperative UAVs guaranteeing stability and safety of the system even in the absence of collision avoidance constraints. Furthermore, the solutions found in this paper may be scaled to accommodate larger teams of UAVs in more complex environments.

As future work on this topic, we will examine how to improve the computational time for the designed controller that allowing the implementation in real time on-board computers in Qball-X4 quadrotors. Also, we will start to apply the designed control policy on a group of UAVs cooperating to perform various tactics.

References

[1] A. Marasco, S. N. Givigi, and C. A. Rabbath, "Dynamic encirclement of a moving target using decentralized nonlinear model predictive control," *American Control Conference*, pp. 3966 – 3972, 2013.

[2] C. Reynolds, "An evolved, vision-based behavioral model of coordinated group motion," *From animals to animats*, vol. 2, pp. 384–392, 1993.

[3] A. Ryan, M. Zennaro, A. Howell, R. Sengupta, and J. Hedrick, "An overview of emerging results in cooperative UAV control," in *43rd IEEE Conference on Decision and Control (CDC)*, vol. 1. IEEE, 2004, pp. 602–607.

[4] Y. Ding, C. Wei, and S. Bao, "Decentralized formation control for multiple UAVs based on leader-following consensus with time-varying delays," in *2013 Chinese Automation Congress (CAC)*. IEEE, 2013, pp. 426–431.

[5] W. Ren and R. Beard, "A decentralized scheme for spacecraft formation flying via the virtual structure approach," in *American Control Conference, 2003. Proceedings of the 2003*, vol. 2. IEEE, 2003, pp. 1746–1751.

[6] J. Lawton, R. Beard, and B. Young, "A decentralized approach to formation maneuvers," *IEEE Transactions on Robotics and Automation*, vol. 19, no. 6, pp. 933–941, 2003.

[7] H. Rezaee, F. Abdollahi, and M. B. Menhaj, "Model-free fuzzy leader-follower formation control of fixed wing UAVs," in *2013 13th Iranian Conference on Fuzzy Systems (IFSC)*. IEEE, 2013, pp. 1–5.

[8] D. Luo, T. Zhou, and S. Wu, "Obstacle avoidance and formation regrouping strategy and control for UAV formation flight," in *10th IEEE International Conference on Control and Automation (ICCA), 2013*. IEEE, 2013, pp. 1921–1926.

[9] A. S. Brandao, J. Barbosa, V. Mendoza, M. Sarcinelli-Filho, and R. Carelli, "A multi-layer control scheme for a centralized UAV formation," in *2014 International Conference on Unmanned Aircraft Systems (ICUAS),*.

[10] G. Regula and B. Lantos, "Formation control of a large group of UAVs with safe path planning," in *21st Mediterranean Conference on Control & Automation (MED)*. IEEE, 2013, pp. 987–993.

[11] E. F. Camacho and C. Bordons, *Model Predictive Control*. London: Springer-Verlag, 2007.

[12] X. Wang, V. Yadav, and S. Balakrishnan, "Cooperative UAV formation flying with obstacle/collision avoidance," *IEEE Transactions on Control Systems Technology,*.

[13] J. Lavaei, A. Momeni, and A. G. Aghdam, "A model predictive decentralized control scheme with reduced communication requirement for spacecraft formation," *IEEE Transactions on Control Systems Technology*, vol. 16, no. 2, pp. 268–278, 2008.

[14] Z. Chao, S.-L. Zhou, L. Ming, and W.-G. Zhang, "UAV formation flight based on nonlinear model predictive control," *Mathematical Problems in Engineering*, vol. 2012, 2012.

[15] F. Allgöwer and A. Zheng, *Nonlinear model predictive control*. Birkhäuser, 2000, vol. 26.

[16] H. Shin, M. Thak, and H. Kim, "Nonlinear model predictive control for multiple UAVs formation using passive sensing," *International Journal of Aeronautical and Space Sciences*, vol. 12, pp. 16–23, 2011.

[17] A. Aswani, P. Bouffard, and C. Tomlin, "Extensions of learning-based model predictive control for real-time application to a quadrotor helicopter," in *American Control Conference (ACC), 2012*. IEEE, 2012, pp. 4661–4666.

[18] P. Bouffard, A. Aswani, and C. Tomlin, "Learning-based model predictive control on a quadrotor: Onboard implementation and experimental results," in *IEEE International Conference on Robotics and Automation (ICRA), 2012*. IEEE, 2012, pp. 279–284.

[19] A. Hafez, S. Givigi, and S. Yousfei, "Unmanned aerial vehicles formation using learning based model predictive control," *Submitted to IEEE Transactions on Aerospace and Electronic Systems*, 2015.

[20] I. Kolmanovsky and E. G. Gilbert, "Theory and computation of disturbance invariant sets for discrete-time linear systems," *Mathematical Problems in Engineering*, vol. 4, no. 4, pp. 317–367, 1998.

[21] A. Aswani, H. Gonzalez, S. S. Sastry, and C. Tomlin, "Provably safe and robust learning-based model predictive control," *Automatica*, vol. 49, no. 5, pp. 1216–1226, 2013.

[22] E. C. Kerrigan, "Robust constraint satisfaction: Invariant sets and predictive control," Ph.D. dissertation, PhD thesis, Cambridge, 2000.

[23] R. Sharma, M. Kothari, C. Taylor, and I. Postlethwaite, "Cooperative target-capturing with inaccurate target information," in *American Control Conference (ACC), 2010*. IEEE, 2010, pp. 5520–5525.

[24] W. B. Dunbar and R. M. Murray, "Distributed receding horizon control for multi-vehicle formation stabilization," *Automatica*, vol. 42, no. 4, pp. 549–558, 2006.

[25] L. Chisci, J. A. Rossiter, and G. Zappa, "Systems with persistent disturbances: predictive control with restricted constraints," *Automatica*, vol. 37, no. 7, pp. 1019–1028, 2001.

[26] E. A. Wan and A. T. Nelson, "Dual extended kalman filter methods," *Kalman filtering and neural networks*, pp. 123–173, 2001.

[27] T. A. Wenzel, K. Burnham, M. Blundell, and R. Williams, "Dual extended kalman filter for vehicle state and parameter estimation," *Vehicle System Dynamics*, vol. 44, no. 2, pp. 153–171, 2006.