Cooperative Filtering and Parameter Estimation for Polynomial PDEs using a Mobile Sensor Network

Ziqiao Zhang, Wencen Wu, and Fumin Zhang

***Abstract*— In this paper, a constrained cooperative Kalman filter is developed to estimate field values and gradients along trajectories of mobile robots collecting measurements. We assume the underlying field is generated by a polynomial partial differential equation with unknown time-varying parameters. A long short-term memory (LSTM) based Kalman filter, is applied for the parameter estimation leveraging the updated state estimates from the constrained cooperative Kalman filter. Convergence for the constrained cooperative Kalman filter has been justified. Simulation results in a 2-dimensional field are provided to validate the proposed method.**

I. Introduction

Many spatial-temporal fields can be described by partial differential equations (PDEs). However, it is often difficult to obtain explicit analytic solution for these PDEs [1]. We need to estimate the field and gradient information in applications such as modeling congested freeway traffic [2], monitoring Arctic sea ice [3], and tracking dynamic pollutant plume propagation [4]. Mobile sensor networks are promising for data collection [5]–[10] that can be used to estimate the state of the PDEs.

In earlier literature, static sensors have been applied to explore PDE fields and the parameters are identified by solving inverse problems of PDEs [11]. The inverse problem, however, can be difficult to solve. Recently, model predictive control [12], adaptive observer [13] and adaptive boundary control [14] have also been studied for parameter identification for PDE models. The two-stage method can be applied to this problem as well [15], with the first stage of estimating field value and gradient and the second stage of identifying PDE parameters using least square method. In our previous works [8], [10], [16], constrained cooperative Kalman filter has been incorporated for state estimation and the recursive least square method has been applied to iteratively update the estimate of unknown parameters.

While our previous works mainly consider the advection-diffusion PDEs with constant coefficient, this paper addresses more general PDE models and unknown time-varying parameters. Similar to our previous works [8], [10], [16], a bootstrap structure of state filter and parameter estimation algorithm is proposed. For the state estimation, the information dynamics are derived, and a constrained cooperative Kalman filter has been proposed. The major difference from our previous works is that the PDE is treated as constraint on the states of the information dynamics. This new structure allows the constrained cooperative Kalman filter to be applied to a class of polynomial PDEs, not just the advection-diffusion PDEs. Another contribution is that an LSTM based Kalman filter revised from our previous work [17] is applied to estimate the parameters of the PDEs, fusing the latest state estimate from the constrained Kalman filter with previous parameter estimates. According to [18], [19], we can prove convergence for the constrained cooperative Kalman filter by showing the unconstrained filter is convergent. Since the PDEs are treated as constraints, the convergence analysis is decoupled from the parameter estimation since the estimated parameters does not appear in the unconstrained filter. These new contributions provide a more general framework for cooperative filtering and parameter identification using mobile sensor networks.

The problem of state estimation and parameter estimation is formulated in Section II. Information dynamics and measurement equation are derived in Section III and Section IV, respectively. Section V introduces the bootstrap structure of state estimation (constrained cooperative Kalman filter) and parameter estimation (LSTM-based Kalman filter). Convergence analysis for state estimation is presented in Section VI. Simulation results are illustrated in Section VII. Conclusions and future work follow in Section VIII.

II. Problem formulation

In this section, we formulate the estimation problem of field parameter and field information along trajectory using mobile sensor networks for a spatial-temporal varying concentration field in d-dimensional space where $d \in \mathbb{N}, d \geq 2$.

We assume that the field can be described by the following partial differential equation in a spatial domain $\Omega \subseteq \mathbb{R}^d$

$$\frac{\partial z(r,t)}{\partial t} = \sum_{i=1}^{M} \theta_i(t)\psi_i(z(r,t), \nabla z(r,t), \nabla^2 z(r,t)), \qquad (1)$$

where $r \in \Omega$ represents location, $t \in \mathbb{R}_+$ represents time, $z(\cdot,\cdot) : \mathbb{R}^d \times \mathbb{R}_+ \to \mathbb{R}$ is the concentration function, ∇ is the gradient operator, and ∇^2 is the Laplacian operator. The partial differential equation is a M-th polynomial function, and $\psi_i(\cdot)$ is the ith order polynomial with time-varying unknown parameter $\theta_i(t)$. One of our goals in this work is to estimate the unknown parameter $\theta_i(t)$.

The research work is supported by ONR grants N00014-19-1-2556 and N00014-19-1-2266; AFOSR grant FA9550-19-1-0283; NSF grants CNS-1828678, S&AS-1849228, GCR-1934836, and CMMI-1917300; NRL grants N00173-17-1-G001 and N00173-19-P-1412; and NOAA grant NA16NOS0120028. Ziqiao Zhang and Fumin Zhang are with the School of Electrical and Computer Engineering, Georgia Institute of Technology, Atlanta, GA 30318, USA. Wencen Wu is with the Computer Engineering Department of San Jose State University, San Jose, CA 95192, USA. Email: `ziqiao.zhang@gatech.edu`, `wencen.wu@sjsu.edu`, `fumin@gatech.edu`

The equation (1) has the initial condition $z(r,0) = z_0(r)$ for $r \in \Omega$, and the boundary condition $z(r,t) = z_b(r,t)$ for $r \in \partial\Omega$, where $z_0(r)$ and $z_b(r,t)$ are arbitrary initial condition and Dirichlet boundary condition, respectively.

Since mobile robots collect discrete measurements at each time step instead of continuous measurements, the continuous PDE model (1) can be discretized using finite difference method at time step t_k as follows

$$\left.\frac{\partial z(r,t)}{\partial t}\right|_{t=t_k} \approx \frac{z(r,k+1)-z(r,k)}{\delta t}, \tag{2}$$

where δt is the sampling interval. By finite difference method (1) will have the following discretization,

$$\begin{aligned} z(r,k+1) &\approx z(r,k) + \delta t \sum_{i=1}^{M} \theta_i(k)\psi_i(z(r,k),\nabla z(r,k),\nabla^2 z(r,k)) \\ &= z(r,k) + \delta t \Psi^\intercal(z(r,k),\nabla z(r,k),\nabla^2 z(r,k))\Theta(k), \end{aligned} \tag{3}$$

where $\Psi(\cdot) = [\psi_1(\cdot),\cdots,\psi_M(\cdot)]^\intercal$ and $\Theta(k) = [\theta_1(k),\theta_2(k),\cdots,\theta_M(k)]^\intercal \in \mathbb{R}^M$.

Suppose a group of N mobile robots are deployed in the field described by (1) and take discrete measurements of the field concentration. Denote the location of agent i at time step t_k as r_i^k and the noisy measurement $p(r_i^k,k)$ taken by this agent at (r_i^k,k) can be written as

$$p(r_i^k,k) = z(r_i^k,k) + n_i, \tag{4}$$

where $n_i \in \mathbb{R}$ is an independently and identically distributed Gaussian noise. Define the center of the mobile robots at t_k as r_c^k and $r_c^k = \frac{1}{N}\sum_{i=1}^{N} r_i^k$. Another goal is to estimate field concentration $z(r_c,t)$ and corresponding gradient $\nabla z(r_c,t)$ along trajectory.

In order to solve this parameter estimation and state estimation problem, we will follow a two-step procedure:

1) We will estimate the field information $z(r_c,t), \nabla z(r_c,t)$ based on the measurements collected by mobile robots.
2) The time-varying unknown parameter $\theta_i(t)$ will be estimated using the estimated states at time t.

III. Information Dynamics

In this section, we will introduce the information dynamics at r_c^k and the general form of discretization for the information dynamics.

First we introduce some simplified notations, $z_{k,k} = z(r_c^k,k)$, $z_{k-1,k-1} = z(r_c^{k-1},k-1)$, $z_{k,k+1} = z(r_c^k,k+1)$. By applying finite difference method, we can get $\frac{\mathrm{d}z}{\mathrm{d}t}$ and $\frac{\partial z}{\partial t}$ as follows,

$$\frac{\mathrm{d}z}{\mathrm{d}t} \approx \frac{z_{k,k}-z_{k-1,k-1}}{\delta t}, \quad \frac{\partial z}{\partial t} \approx \frac{z_{k,k+1}-z_{k,k}}{\delta t}, \tag{5}$$

where δt is the sampling interval. Then we can obtain

$$\begin{aligned} \frac{\mathrm{d}z}{\mathrm{d}t} - \frac{\partial z}{\partial t} &\approx \frac{1}{\delta t}\left(z_{k,k}-z_{k-1,k-1}-z_{k,k+1}+z_{k,k}\right) \\ &= \frac{1}{\delta t}\left(2z_{k,k}-z_{k-1,k-1}-z_{k,k+1}\right). \end{aligned} \tag{6}$$

We also know that

$$\frac{\mathrm{d}z}{\mathrm{d}t} - \frac{\partial z}{\partial t} = \frac{1}{\delta t}(r_c^k - r_c^{k-1})^\intercal \nabla z_{k-1,k-1}. \tag{7}$$

From (6) and (7), we can have

$$\left(2z_{k,k}-z_{k-1,k-1}-z_{k,k+1}\right) = (r_c^k - r_c^{k-1})^\intercal \nabla z_{k-1,k-1}. \tag{8}$$

Since $z_{k,k} = z_{k-1,k} + (r_c^k - r_c^{k-1})^\intercal \nabla z_{k-1,k}$, we should have

$$\begin{aligned} z_{k,k+1} =& 2z_{k,k} - z_{k-1,k-1} - (r_c^k - r_c^{k-1})^\intercal \nabla z_{k-1,k-1} \\ =& 2z_{k,k} - 2(r_c^k - r_c^{k-1})^\intercal \nabla z_{k-1,k} \\ & - z_{k-1,k-1} + (r_c^k - r_c^{k-1})^\intercal \nabla z_{k-1,k-1} \\ & + 2(r_c^k - r_c^{k-1})^\intercal \nabla z_{k-1,k} - 2(r_c^k - r_c^{k-1})^\intercal \nabla z_{k-1,k-1} \\ =& 2z_{k-1,k} - z_{k-1,k-1} + (r_c^k - r_c^{k-1})^\intercal \nabla z_{k-1,k-1} + \text{h.o.t.}. \end{aligned} \tag{9}$$

By ignoring the higher order term, (9) can be written as

$$z_{k,k+1} = 2z_{k-1,k} - z_{k-1,k-1} + (r_c^k - r_c^{k-1})^\intercal \nabla z_{k-1,k-1}. \tag{10}$$

We also have that

$$\frac{\mathrm{d}z}{\mathrm{d}t} \approx \frac{z_{k,k+1}-z_{k-1,k}}{\delta t}, \quad \frac{\partial z}{\partial t} \approx \frac{z_{k,k+1}-z_{k,k}}{\delta t}. \tag{11}$$

Similarly, we can get

$$\frac{\mathrm{d}z}{\mathrm{d}t} - \frac{\partial z}{\partial t} = \frac{1}{\delta t}\left(-z_{k-1,k}+z_{k,k}\right) = \frac{1}{\delta t}(r_c^k - r_c^{k-1})^\intercal \nabla z_{k-1,k}. \tag{12}$$

Thus, we have the following two relationships

$$z_{k,k} = z_{k-1,k} + (r_c^k - r_c^{k-1})^\intercal \nabla z_{k-1,k} \tag{13}$$

$$\begin{aligned} z_{k,k+1} =& 2z_{k-1,k} - 2(r_c^k - r_c^{k-1})^\intercal \nabla z_{k-1,k} - z_{k-1,k-1} \\ & + (r_c^k - r_c^{k-1})^\intercal \nabla z_{k-1,k-1} \end{aligned} \tag{14}$$

Define state variable $x(k) = [z_{k-1,k-1}, \nabla z_{k-1,k-1}, z_{k-1,k}, \nabla z_{k-1,k}]^\intercal$, and we can get the following state equation

$$x(k+1) = A(k)x(k) + U(k) + e(k), \tag{15}$$

where $A(k) \triangleq \begin{bmatrix} 0 & 0 & 1 & (r_c^k - r_c^{k-1})^\intercal \\ 0 & I_{d\times d} & 0 & 0 \\ -1 & (r_c^k - r_c^{k-1})^\intercal & 2 & -2(r_c^k - r_c^{k-1})^\intercal \\ 0 & 0 & 0 & I \end{bmatrix}$, $U(k) \triangleq \begin{bmatrix} 0 \\ \nabla^2 z_{k-1,k-1}(r_c^k - r_c^{k-1}) \\ 0 \\ \nabla^2 z_{k-1,k}(r_c^k - r_c^{k-1}) \end{bmatrix}$ and $e(k)$ is the noise term.

IV. Measurement Equation

The field concentration can be locally approximated by a Taylor series up to second order as

$$\begin{aligned} z(r_i^{k-1},k-1) \approx& z(r_c^{k-1},k-1) + (r_i^{k-1} - r_c^{k-1})^\intercal \nabla z(r_c^{k-1},k-1) \\ & + \frac{1}{2}(r_i^{k-1} - r_c^{k-1})^\intercal H(r_c^{k-1},k-1)(r_i^{k-1} - r_c^{k-1}), \end{aligned} \tag{16}$$

$$\begin{aligned} z(r_i^k,k) \approx& z(r_c^{k-1},k) + (r_i^k - r_c^{k-1})^\intercal \nabla z(r_c^{k-1},k) \\ & + \frac{1}{2}(r_i^k - r_c^{k-1})^\intercal H(r_c^{k-1},k)(r_i^k - r_c^{k-1}). \end{aligned} \tag{17}$$

Let $Z(k) = [z(r_1^{k-1}, k-1), \cdots, z(r_N^{k-1}, k-1), z(r_1^k, k), \cdots, z(r_N^k, k)]^\intercal$ be the vector of true field values. Define the matrices $C(k)$ and $D(k)$ as

$$C(k) = \begin{bmatrix} 1 & (r_1^{k-1} - r_c^{k-1})^\intercal & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots \\ 1 & (r_N^{k-1} - r_c^{k-1})^\intercal & 0 & 0 \\ 0 & 0 & 1 & (r_1^k - r_c^{k-1})^\intercal \\ \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & 1 & (r_N^k - r_c^{k-1})^\intercal \end{bmatrix}, \quad (18)$$

and

$$D(k) = \begin{bmatrix} \frac{1}{2}((r_1^{k-1} - r_c^{k-1}) \otimes (r_1^{k-1} - r_c^{k-1}))^\intercal \\ \vdots \\ \frac{1}{2}((r_N^{k-1} - r_c^{k-1}) \otimes (r_N^{k-1} - r_c^{k-1}))^\intercal \\ \frac{1}{2}((r_1^k - r_c^{k-1}) \otimes (r_1^k - r_c^{k-1}))^\intercal \\ \vdots \\ \frac{1}{2}((r_N^k - r_c^{k-1}) \otimes (r_N^k - r_c^{k-1}))^\intercal \end{bmatrix}, \quad (19)$$

where $\otimes$ is the Kronecker product. The Taylor expansions (16) for all sensors near r_c^k can be rewritten in a vector form as

$$Z(k) = C(k)x(k) + D(k)H(k), \quad (20)$$

where $H(k)$ is a column vector obtained by rearranging elements of the Hessian $H(r_c^{k-1}, k-1)$.

Suppose $\hat{H}(k)$ represents the estimate of the vector form Hessian $H(k)$ at the center r_c^k, equation (4) can be remodeled as

$$P(k) = C(k)x(k) + D(k)\hat{H}(k) + D(k)\varepsilon(k) + n(k), \quad (21)$$

where $P(k) = [p(r_1^{k-1}, k-1), \cdots, p(r_N^{k-1}, k-1), p(r_1^k, k), \cdots, p(r_N^k, k)]^\intercal$ is the measurement vector, $\varepsilon(k)$ represents the error in the estimation of the Hessian matrices, and $\mathbf{n}(k)$ is the vector of Gaussian measurement noise n_i in equation (4). For the estimation of Hessian matrices, we will follow the procedure of cooperative estimation in [20], and we will not discuss the details of how to get $\hat{H}(k)$.

V. Bootstrap Structure of State Estimation and Parameter Estimation

In this section, we will construct a bootstrap algorithm for state estimation and parameter estimation. As shown in Fig. 1, the proposed algorithm is composed of two parts, a constrained cooperative Kalman filter for state estimation and an LSTM-based Kalman filter (LSTM-KF) for parameter estimation. At time step t_k, the constrained cooperative Kalman filter updates the state estimation $\hat{x}(k)$ by incorporating the previous parameter estimation $\hat{\Theta}(k-1)$ from the LSTM-KF. Then the LSTM-KF updates the parameter $\hat{\Theta}(k)$ according to the updated state estimation $\hat{x}(k)$.

A. Constrained Cooperative Kalman Filter

We construct a constrained Kalman filter for state estimation. The idea is to use information dynamics as the state equation, build the connection between state and measurement as measurement equation, and include approximated PDE model as constraint for this system.

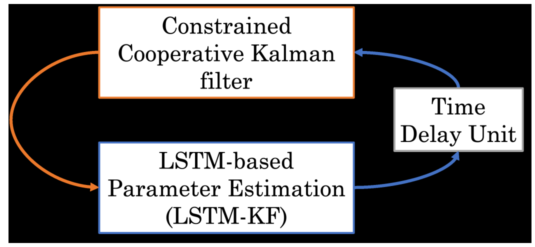

Fig. 1. Bootstrap structure of the proposed algorithm

Before we introduce the constrained cooperative Kalman filter, we need to linearize the PDE constraint (3) around state variable $x(k) = [z_{k-1,k-1}, \nabla z_{k-1,k-1}, z_{k-1,k}, \nabla z_{k-1,k}]^\intercal$. At $x(k)$, (3) can be rewritten as follows,

$$F(k) \triangleq z_{k-1,k} - z_{k-1,k-1} - \delta t \Psi(z_{k-1,k-1}, \nabla z_{k-1,k-1}, \nabla^2 z_{k-1,k-1})^\intercal \hat{\Theta}(k-1) = 0, \quad (22)$$

where $z_{k-1,k-1}, \nabla z_{k-1,k-1}, \nabla^2 z_{k-1,k-1}, z_{k-1,k}$ are treated as independent variables during linearization. After taking partial derivative of left hand side of (22) w.r.t. $x(k)$, we can obtain

$$J(k) \triangleq \begin{bmatrix} \frac{\partial}{\partial z_{k-1,k-1}} F(k) \\ \frac{\partial}{\partial \nabla z_{k-1,k-1}} F(k) \\ \frac{\partial}{\partial z_{k-1,k}} F(k) \\ \frac{\partial}{\partial \nabla z_{k-1,k}} F(k) \end{bmatrix}^\intercal = \begin{bmatrix} \frac{\partial}{\partial z_{k-1,k-1}} F(k) \\ \frac{\partial}{\partial \nabla z_{k-1,k-1}} F(k) \\ 1 \\ 0 \end{bmatrix}^\intercal. \quad (23)$$

And the linearized model at $x(k)$ can be written as

$$J(k)(x - x(k)) - F(k) = 0. \quad (24)$$

Assumption V.1 *We assume that the noises $e(k), \varepsilon(k), n(k)$ are i.i.d. Gaussian noise with zero mean, and with constant covariance matrix, i.e. $E[e(k)e^\intercal(k)] = W$, $E[n(k)n^\intercal(k)] = R$ and $E[\varepsilon(k)\varepsilon^\intercal(k)] = Q$.*

The constrained cooperative Kalman filter can be constructed using 6 steps:

(1) One-step state prediction

$$\hat{x}^-(k) = A(k-1)\tilde{x}^+(k-1) + U(k-1), \quad (25)$$

where $\tilde{x}^+(k-1)$ is the constrained state estimate from previous time step, and $\hat{x}^-(k)$ is the one-step state prediction.

(2) Error covariance of $\hat{x}^-(k)$

$$R^-(k) = A(k-1)R^+(k-1)A^\intercal(k-1) + W. \quad (26)$$

(3) Optimal gain

$$K(k) = R^-(k)C^\intercal(k)[C(k)R^-(k)C^\intercal(k) + D(k)QD^\intercal(k) + R]^{-1}. \quad (27)$$

(4) Updated unconstrained state estimate

$$\hat{x}^+(k) = \hat{x}^-(k) + K(k)(P(k) - C(k)\hat{x}^-(k) - D(k)\hat{H}(k)). \quad (28)$$

(5) Error covariance of $\hat{x}^+(k)$

$$(R^+(k))^{-1} = (R^-(k))^{-1} + C^\intercal(k)[D(k)QD^\intercal(k) + R]^{-1}C(k). \quad (29)$$

(6) Updated constrained state estimate

$$\tilde{x}^+(k) = \hat{x}^+(k) + J(k)^\intercal [J(k) J^\intercal(k)]^{-1} \hat{F}(k), \tag{30}$$

where $\hat{x}^+(k) = \left[\hat{z}^+_{k-1,k-1}, \nabla\hat{z}^+_{k-1,k-1}, \hat{z}^+_{k-1,k}, \nabla\hat{z}^+_{k-1,k}\right]^\intercal$ and $\hat{F}(k) = \hat{z}^+_{k-1,k} - \hat{z}^+_{k-1,k-1} - \delta t \Psi(\hat{z}^+_{k-1,k-1}, \nabla\hat{z}^+_{k-1,k-1}, \nabla^2\hat{z}^+_{k-1,k-1})^\intercal \hat{\Theta}(k-1)$. This is because $J(k)(\tilde{x}^+(k) - \hat{x}^+(k)) - \hat{F}(k) = 0$ according to (24).

B. Parameter Estimation

Since the unknown parameter $\Theta(k)$ is time-varying, recursive least square method [16] cannot be applied to estimate non-constant parameters. If there is no plenty of history data, random walk model [21] may be applied for parameter estimation. In our case, we are considering the scenario when there is abundant history data of estimated parameters. For example, models for underwater and atmospheric environments are usually complicated and require heavy computation complexity to be run online [22], [23]. There exist cloud computing models [24], [25] for data assimilation in these environments, and such models contain large amounts of history data which can be used for training purposes and initialization of neural networks.

Long short-term memory (LSTM) is a variant of recurrent neural networks, and it can build both long-term dependencies and short-term relationships of sequential data [26], [27]. The time-varying unknown parameter $\theta_i(k)$ can be viewed as this type of sequential data and we consider using the LSTM to predict the unknown $\theta_i(k)$ using historical data from cloud computing model.

Assume the dynamics of $\Theta(k)$ can be described by the following equation

$$\Theta(k) = g(\Theta(k-1), \Theta(k-2), \cdots, \Theta(k-L)), \tag{31}$$

where $g(\cdot)$ is an unknown function and constant $L \in \mathbb{N}_+$ describes the time dependency of parameter Θ.

By applying LSTM to approximate function $g(\cdot)$, we can obtain the following relationship,

$$\hat{\Theta}(k) = g_{LSTM}(\hat{\Theta}(k-1), \hat{\Theta}(k-2), \cdots, \hat{\Theta}(k-L)). \tag{32}$$

Next we will apply an LSTM-based Kalman filter from our previous work [17] to update the LSTM prediction by incorporating the estimated $z, \nabla z, \nabla^2 z$.

$$\begin{aligned}
\hat{\Theta}^-(k) &= g_{LSTM}(\hat{\Theta}^+(k-1), \hat{\Theta}^+(k-2), \cdots, \hat{\Theta}^+(k-L)),\\
P^-_\theta(k) &= G(k-1) P^+_\theta(k-1) G(k-1)^\intercal + W_\theta(k-1),\\
K_\theta(k) &= P^-_\theta(k) \hat{\Psi}(k) [\hat{\Psi}^\intercal(k) P^-_\theta(k) \hat{\Psi}(k) + R_\theta(k)]^{-1},\\
\hat{\Theta}^+(k) &= \hat{\Theta}^-(k) + K_\theta(k)(p_\theta(k) - \hat{\Psi}^\intercal(k)\hat{\Theta}^-(k)),\\
P^+_\theta(k)^{-1} &= P^-_\theta(k)^{-1} + \hat{\Psi}(k) R_\theta(k)^{-1} \hat{\Psi}^\intercal(k),
\end{aligned} \tag{33}$$

where $p_\theta(k) = (\hat{z}_{k-1,k} - \hat{z}_{k-1,k-1})/\delta t$, W_θ is the covariance for state estimation error, and R_θ is the covariance for measurement noise. The matrix $G(k-1)$ is defined as $G(k-1) = [I_{M\times M},\ 0_{M\times(LM)}]$. The vector $\hat{\Psi}(k)$ is defined as $\hat{\Psi}(k) = \Psi(\hat{z}_{k-1,k}, \nabla\hat{z}_{k-1,k}, \nabla^2\hat{z}_{k-1,k})$, and it is a polynomial vector that contains all the polynomials $\{\psi_i(\cdot)\}_{i=1}^M$ in (1) at $(\hat{z}_{k-1,k}, \nabla\hat{z}_{k-1,k}, \nabla^2\hat{z}_{k-1,k})$. Field estimate $\hat{z}_{k-1,k}$ and gradient estimate $\nabla\hat{z}_{k-1,k}$ can be obtained from state estimate $\hat{x}(k)$ and $\nabla^2\hat{z}_{k-1,k}$ can be estimated according to Hessian estimation.

VI. Convergence Analysis

In this section, we will study the sufficient conditions for the convergence of the unconstrained system. The analysis is composed of two parts: uniformly complete controllability and uniformly complete observability. For the unconstrained system, we do not need the parameter Θ, which means that the convergence analysis for the unconstrained system is decoupled from the parameter estimation. If the unconstrained Kalman filter is convergent, then the convergence of the constrained Kalman filter can be guaranteed [18], [19].

Let $\Phi(k,j)$ be the state transition matrix from time t_j to t_k, where $k > j$. Then, $\Phi(k,j) = A(k-1)A(k-2)\cdots A(j) = \Phi^{-1}(j,k)$, and we can have the following lemma.

Lemma VI.1 *For $\Phi(k,j)$ as defined above and $C(k)$ as defined in (18), we can have*

$$\Phi(k,j) = \begin{bmatrix} j-k+1 & \phi_{12} & k-j & \phi_{14} \\ 0 & I_{d\times d} & 0 & 0 \\ j-k & \phi_{32} & k-j+1 & \phi_{34} \\ 0 & 0 & 0 & I \end{bmatrix}, \tag{34}$$

$$\Phi(j,k) = \begin{bmatrix} 1+k-j & \phi'_{12} & j-k & \phi'_{14} \\ 0 & I_{d\times d} & 0 & 0 \\ k-j & \phi'_{32} & j-k+1 & \phi'_{34} \\ 0 & 0 & 0 & I \end{bmatrix}, \tag{35}$$

where

$$\phi_{12} = \sum_{l=j+1}^{k-1} (k-l)(r_c^{l-1} - r_c^{l-2})^\intercal,$$

$$\phi_{14} = (r_c^{k-1} - r_c^{k-2})^\intercal - \sum_{l=j+1}^{k-2} (k-l-1)(r_c^{l-1} - r_c^{l-2})^\intercal,$$

$$\phi_{32} = \sum_{l=j+1}^{k} (k-l+1)(r_c^{l-1} - r_c^{l-2})^\intercal,$$

$$\phi_{34} = -\sum_{l=j+1}^{k-1} (k-l)(r_c^{l-1} - r_c^{l-2})^\intercal,$$

$$\begin{aligned}
\phi'_{12} &= -(k-j+1)\phi_{12} - (j-k)\phi_{32},\\
\phi'_{14} &= -(k-j+1)\phi_{14} - (j-k)\phi_{34},\\
\phi'_{32} &= -(k-j)\phi_{12} - (j-k+1)\phi_{32},\\
\phi'_{34} &= -(k-j)\phi_{14} - (j-k+1)\phi_{34},
\end{aligned}$$

and

$$C(j)\Phi(j,k) = \begin{bmatrix} (C\Phi)_{11} & (C\Phi)_{12} \\ (C\Phi)_{21} & (C\Phi)_{22} \end{bmatrix}, \tag{36}$$

where $(C\Phi)_{11} = \begin{bmatrix} 1+k-j & \phi'_{12} + (r_1^{j-1} - r_c^{j-1})^\intercal \\ \vdots & \vdots \\ 1+k-j & \phi'_{12} + (r_N^{j-1} - r_c^{j-1})^\intercal \end{bmatrix}$, $(C\Phi)_{12} = \begin{bmatrix} j-k & \phi'_{14} \\ \vdots & \vdots \\ j-k & \phi'_{14} \end{bmatrix}$, $(C\Phi)_{21} = \begin{bmatrix} k-j & \phi'_{32} \\ \vdots & \vdots \\ k-j & \phi'_{32} \end{bmatrix}$,

$$(C\Phi)_{22} = \begin{bmatrix} j-k+1 & \phi'_{34} + (r_1^j - r_c^{j-1})^\intercal \\ \vdots & \vdots \\ j-k+1 & \phi'_{34} + (r_N^j - r_c^{j-1})^\intercal \end{bmatrix}.$$

Lemma VI.2 *For a matrix $B \in \mathbb{R}^{m\times n}$, if each element of B is bounded, then each element of $BB^\intercal$ and $B^\intercal B$ is also bounded.*

A. Uniformly Complete Controllability

Definition VI.3 *The proposed unconstrained cooperative filter is uniformly completely controllable if there exist $\tau_1 > 0$, $\lambda_1 > 0$, and $\lambda_2 > 0$ such that the controllability Grammian $\mathfrak{C}(k,k-\tau_1) = \sum_{j=k-\tau_1}^{k} \Phi(k,j) W \Phi(k,j)^\intercal$ satisfies $\lambda_1 I_{2(d+1)\times 2(d+1)} \le \mathfrak{C}(k,k-\tau_1) \le \lambda_2 I_{2(d+1)\times 2(d+1)}$ for all $k > \tau_1$.*

Proposition VI.4 *The proposed unconstrained cooperative filter is uniformly completely controllable if the following conditions are satisfied:*

(Cd1) The covariance matrix W is bounded, i.e., $\lambda_3 I \le W \le \lambda_4 I$ for some constants $\lambda_3, \lambda_4 > 0$.

(Cd2) The speed of each agent is uniformly bounded, i.e., $\|r_i^j - r_i^{j-1}\| \le \lambda_5$ for all time j, for $i = 1,\cdots,N$, and for some constant $\lambda_5 > 0$.

Proof: Based on (Cd1), the controllability Grammian satisfies $\lambda_3 \sum_{j=k-\tau_1}^{k} \Phi(k,j)\Phi^\intercal(k,j) \le \mathfrak{C}(k,k-\tau_1) \le \lambda_4 \sum_{j=k-\tau_1}^{k} \Phi(k,j)\Phi^\intercal(k,j)$ for any k, τ_1 such that $k > \tau_1$. If we can find uniform bounds for $\Phi(k,j)\Phi^\intercal(k,j)$, then the upper and lower bounds for the controllabiligy Grammian can also be found.

First we will show the existence of an upper bound. According to (Cd2), the agents' speed is bounded by λ_5, which implies the speed of formation center r_c is also bounded since for any l

$$\begin{aligned} \|r_c^l - r_c^{l-1}\| &= \|\frac{1}{N}\sum_{i=1}^{N} r_i^l - \frac{1}{N}\sum_{i=1}^{N} r_i^{l-1}\| = \frac{1}{N}\|\sum_{i=1}^{N}(r_i^l - r_i^{l-1})\| \\ &\le \frac{1}{N}\sum_{i=1}^{N}\|r_i^l - r_i^{l-1}\| \le \lambda_5. \end{aligned} \tag{37}$$

This implies that each element of $\Phi(k,j)$ is bounded and each element of $\Phi(k,j)\Phi^\intercal(k,j)$ is also bounded by Lemma VI.2. Therefore, there exists $\lambda_6 > 0$ such that $\Phi(k,j)\Phi^\intercal(k,j) \le \lambda_6 I_{2(d+1)\times 2(d+1)}$ for $j \in [k-\tau_1, k]$. Then by defining $\lambda_2 = \lambda_4 \tau_1 \lambda_6 > 0$ we can get $\mathfrak{C}(k,k-\tau_1) \le \lambda_4 \sum_{j=k-\tau_1}^{k} \Phi(k,j)\Phi^\intercal(k,j) \le \lambda_4 \tau_1 \lambda_6 I_{2(d+1)\times 2(d+1)} = \lambda_2 I_{2(d+1)\times 2(d+1)}$.

Next we want to derive the positive lower bound for $\Phi(k,j)\Phi^\intercal(k,j)$. Since it is hard to derive the explicit form of the eigenvalues of $\Phi(k,j)\Phi^\intercal(k,j)$, we want to show the eigenvalues are positive.

We can show $\det(A(l)) = 1$ for any l as follows,

$$\begin{aligned} \det(A(l)) &= \det\left(\begin{bmatrix} 0 & 0 & 1 & (r_c^l - r_c^{l-1})^\intercal \\ 0 & I_{d\times d} & 0 & 0 \\ -1 & (r_c^l - r_c^{l-1})^\intercal & 2 & 0 \\ 0 & 0 & 0 & I_{d\times d} \end{bmatrix}\right) \\ &= \det\left(\begin{bmatrix} 0 & 0 & 1 \\ 0 & I_{d\times d} & 0 \\ -1 & (r_c^l - r_c^{l-1})^\intercal & 2 \end{bmatrix}\right) \\ &= -\det\left(\begin{bmatrix} 1 & 0 & 0 \\ 0 & I_{d\times d} & 0 \\ 2 & (r_c^l - r_c^{l-1})^\intercal & -1 \end{bmatrix}\right) = 1. \end{aligned}$$

Then according to the definition of $\Phi(k,j)$, we can have

$$\begin{aligned} \det(\Phi(k,j)) &= \det(A(k-1)A(k-2)\cdots A(j)) \\ &= \prod_{l=j}^{k-1} \det(A(l)) = 1. \end{aligned}$$

This means that

$$\det(\Phi(k,j)\Phi^\intercal(k,j)) = \det(\Phi(k,j)(\det(\Phi(k,j)^\intercal) = 1\times 1 = 1.$$

Thus, we can obtain

$$\prod_{i=1}^{6} \mathrm{eig}_i(\Phi(k,j)\Phi^\intercal(k,j)) = \det(\Phi(k,j)\Phi^\intercal(k,j)) = 1, \tag{38}$$

which implies the eigenvalues are all non-zero value. Since $\Phi(k,j)\Phi^\intercal(k,j)$ is real, symmetric and positive semi-definite, this means the eigenvalues are all positive. Hence, there exists $\lambda_7 > 0$ such that $\lambda_7 I_{2(d+1)\times 2(d+1)} \le \Phi(k,j)\Phi^\intercal(k,j)$ for all $j \in [k-\tau_1, k]$. Then by defining $\lambda_1 = \lambda_3 \tau_1 \lambda_7 > 0$ we can have $\mathfrak{C}(k,k-\tau_1) \ge \lambda_3 \sum_{j=k-\tau_1}^{k} \Phi(k,j)\Phi^\intercal(k,j) \le \lambda_3 \tau_1 \lambda_7 I_{2(d+1)\times 2(d+1)} = \lambda_1 I_{2(d+1)\times 2(d+1)}$.

Therefore, there exist $\lambda_1, \lambda_2 > 0$ such that $\lambda_1 I_{2(d+1)\times 2(d+1)} \le \mathfrak{C}(k,k-\tau_1) \le \lambda_2 I_{2(d+1)\times 2(d+1)}$ for all $k > \tau_1$ and the uniformly complete controllability has been proved. ■

B. Uniformly Complete Observability

In this part, we analyze the observability of the state dynamics (15) with the measurement (21).

For uniformly complete observability, the following sufficient conditions are established for a moving formation.

Definition VI.5 *The proposed unconstrained cooperative filter is uniformly completely observable if there exist $\tau_2 > 0$, $\lambda_8 > 0$, and $\lambda_9 > 0$ such that the observability Grammian $\mathfrak{O}(k,k-\tau_2) = \sum_{j=k-\tau_2}^{k} \Phi^\intercal(j,k) C^\intercal(j) [D(j) Q D^\intercal(j) + R]^{-1} C(j) \Phi(j,k)$ satisfies $\lambda_8 I_{2(d+1)\times 2(d+1)} \le \mathfrak{O}(k,k-\tau_2) \le \lambda_9 I_{2(d+1)\times 2(d+1)}$ for all $k > \tau_2$.*

Proposition VI.6 *The proposed unconstrained cooperative filter is uniformly completely observable if (Cd2) and the following conditions are satisfied:*

(Cd3) The number of agents N satisfies $N > d$.

(Cd4) The covariance matrices R and Q are bounded, i.e., $\lambda_{10} I \le R \le \lambda_{11} I$ and $0 \le Q \le \lambda_{12} I$ for some constants $\lambda_{10}, \lambda_{11}, \lambda_{12} > 0$.

(Cd5) The distance between each agent and the formation center is uniformly bounded from both above and below, i.e., $\lambda_{13} \leq \|r_i^{j-1} - r_c^{j-1}\| \leq \lambda_{14}$ for all j, for $i = 1,2,...,N$, and for some constants $\lambda_{13}, \lambda_{14} > 0$.

(Cd6) There exists a constant time difference τ_2, and for all $k > \tau_2$, there exists a time instance $j_1 \in [k-\tau_2, k]$, as well as two sets of agents indexed by $\{i_1, \cdots, i_d\}, \{i_{d+1}, \cdots, i_{2d}\}$ respectively, such that $(r_{i_1}^{j_1-1} - r_c^{j_1-1}), \cdots, (r_{i_d}^{j_1-1} - r_c^{j_1-1})$ are linearly independent, and $(r_{i_{d+1}}^{j_1} - r_c^{j_1-1}), \cdots, (r_{i_{2d}}^{j_1} - r_c^{j_1-1})$ are linearly independent.

Proof:

Based on (Cd5), each element in $D(j)$ is bounded. According to (Cd5), there exist $\lambda_{15}, \lambda_{16} > 0$ such that $\lambda_{15} I \leq D(j) Q D^\intercal(j) + R \leq \lambda_{16} I$ for any $j \in [k-\tau_2, k]$. Then the observability Grammian satisfies $\lambda_{16}^{-1} \sum_{j=k-\tau_2}^{k} \Phi^\intercal(j,k) C^\intercal(j) C(j) \Phi(j,k) \leq \mathfrak{O}(k, k-\tau_2) \leq \lambda_{15}^{-1} \sum_{j=k-\tau_2}^{k} \Phi^\intercal(j,k) C^\intercal(j) C(j) \Phi(j,k)$ for any k, τ_2 such that $k > \tau_2$.

From (Cd2) and (Cd6), we can know that each element of $C(j)\Phi(j,k)$ is bounded, which implies that each element of $\Phi^\intercal(j,k) C^\intercal(j) C(j) \Phi(j,k)$ is also bounded by Lemma VI.2. Therefore, there exists $\lambda_{17} > 0$ such that $\Phi^\intercal(j,k) C^\intercal(j) C(j) \Phi(j,k) \leq \lambda_{17} I_{2(d+1)\times 2(d+1)}$ for $j \in [k-\tau_2, k]$. Then by defining $\lambda_9 = \lambda_{15}^{-1} \tau_2 \lambda_{17} > 0$ we can get $\mathfrak{O}(k, k-\tau_2) \leq \lambda_{15}^{-1} \sum_{j=k-\tau_2}^{k} \Phi^\intercal(j,k) C^\intercal(j) C(j) \Phi(j,k) \leq \lambda_{15}^{-1} \tau_2 \lambda_{17} I_{2(d+1)\times 2(d+1)} = \lambda_9 I_{2(d+1)\times 2(d+1)}$.

According to Lemma VI.1, the matrix $C(j)\Phi(j,k)$ can be simplified using elementary column operations as follows

$$C(j)\Phi(j,k) \to \begin{bmatrix} 1 & (r_1^{j-1} - r_c^{j-1})^\intercal & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots \\ 1 & (r_N^{j-1} - r_c^{j-1})^\intercal & 0 & 0 \\ 0 & 0 & 1 & (r_1^{j} - r_c^{j-1})^\intercal \\ \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & 1 & (r_N^{j} - r_c^{j-1})^\intercal \end{bmatrix} \triangleq (C\Phi)'. \quad (39)$$

Consider the two sets of agents $\{i_1, \cdots, i_d\}, \{i_{d+1}, \cdots, i_{2d}\} \subseteq \{1, \cdots, N\}$ given by (Cd6). Since $(r_{i_1}^{j_1-1} - r_c^{j_1-1})$, $\cdots, (r_{i_d}^{j_1-1} - r_c^{j_1-1})$ are linearly independent. This means that the matrix $\begin{bmatrix} (r_1^{j_1-1} - r_c^{j_1-1})^\intercal \\ \vdots \\ (r_N^{j_1-1} - r_c^{j_1-1})^\intercal \end{bmatrix} \in \mathbb{R}^{N\times d}$ is composed of d linearly independent column vectors, and the d column vectors are independent of $\begin{bmatrix} 1 \\ \vdots \\ 1 \end{bmatrix} \in \mathbb{R}^{N\times 1}$. Thus, the matrix $\begin{bmatrix} 1 & (r_1^{j_1-1} - r_c^{j_1-1})^\intercal \\ \vdots & \vdots \\ 1 & (r_N^{j_1-1} - r_c^{j_1-1})^\intercal \end{bmatrix} \in \mathbb{R}^{N\times (d+1)}$ is composed of $(d+1)$ linearly independent column vectors.

We can use similar approach to show the matrix $\begin{bmatrix} 1 & (r_1^{j_1} - r_c^{j_1-1})^\intercal \\ \vdots & \vdots \\ 1 & (r_N^{j_1} - r_c^{j_1-1})^\intercal \end{bmatrix} \in \mathbb{R}^{N\times (d+1)}$ is composed of $(d+1)$ linearly independent column vectors. Thus, the matrix $(C\Phi)'$ has full column rank.

Since column operations does not change the column rank of the original matrix, the matrix $C(j_1)\Phi(j_1,k)$ also has full column rank. This means that the matrix $\Phi^\intercal(j_1,k) C^\intercal(j_1) C(j_1) \Phi(j_1,k)$ is invertible and has full rank, which implies $\Phi^\intercal(j_1,k) C^\intercal(j_1) C(j_1) \Phi(j_1,k)$ is positive definite. Then the summation matrix $\sum_{j=k-\tau_2}^{k} \Phi^\intercal(j,k) C^\intercal(j) C(j) \Phi(j,k)$ is strictly positive definite. Hence, there exists $\lambda_{18} > 0$ such that $\lambda_{18} I_{2(d+1)\times 2(d+1)} \leq \sum_{j=k-\tau_2}^{k} \Phi^\intercal(j,k) C^\intercal(j) C(j) \Phi(j,k)$ for all $k > \tau_2$. Then by defining $\lambda_8 = \lambda_{18} \lambda_{16}^{-1} > 0$ we can have $\mathfrak{O}(k, k-\tau_2) \geq \lambda_{18} \sum_{j=k-\tau_2}^{k} \Phi^\intercal(j,k) C^\intercal(j) C(j) \Phi(j,k) \geq \lambda_{18} \lambda_{16}^{-1} I_{2(d+1)\times 2(d+1)} = \lambda_8 I_{2(d+1)\times 2(d+1)}$.

Therefore, there exist $\lambda_8, \lambda_9 > 0$ such that $\lambda_8 I_{2(d+1)\times 2(d+1)} \leq \mathfrak{O}(k, k-\tau_2) \leq \lambda_9 I_{2(d+1)\times 2(d+1)}$ for all $k > \tau_2$ and the uniformly complete observability has been proved.

■

According to Theorem 7.4 in [28], the convergence of the unconstrained cooperative Kalman filter can be guaranteed by uniformly complete controllability and observability.

In [18], [19], convergence for Kalman filter with linear and nonlinear state equality constraint has been studied. Similar approach can be applied to the convergence analysis of our constrained filter. In (30), the constrained state estimate $\tilde{x}^+$ can be viewed as the projection of unconstrained estimate $\hat{x}^+$ onto the space of constraint $\hat{F} = 0$. Denote x as the true state value. Then according to Theorem 4 in [18],

$$\|x - \tilde{x}^+\|_2 \leq \|x - \hat{x}^+\|_2, \quad (40)$$

where $\|\cdot\|_2$ is the l_2 norm. Since the unconstrained cooperative Kalman filter is convergent, the constrained cooperative Kalman filter is also convergent.

Remark VI.7 *The estimated parameter $\hat{\Theta}$ only appears in the PDE constraint and is not included in either the state equation or the measurement equation. Therefore, the way we setup the information dynamics helps separate the state estimation and parameter estimation, and makes it easy to generalize to different PDE models with different parameters that need to be estimated.*

VII. Simulation Result

`Improve the performance of parameter estimation.`

In this section, we present a simulation in $\mathbb{R}^2$ that demonstrates the proposed algorithm enables the mobile sensor network to estimate state and parameter along trajectory.

A group of four agents with a square-shape formation (shown in Fig.2) are employed in the PDE field for 600 time steps with sampling interval $\delta t = 0.1s$.

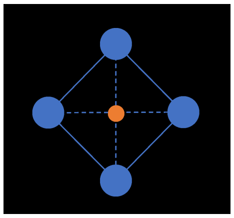

Fig. 2. The formation of the mobile sensor network consisting of four agents.

The four agents have the same velocity $v(k)$, which is predetermined as follows

$$v(k)=\begin{cases}[1,\ 0]^{\intercal}, & 0\leq k\leq 150,\\ [0,\ 1]^{\intercal}, & 151\leq k\leq 300,\\ [-1,\ 0]^{\intercal}, & 301\leq k\leq 450,\\ [0,\ -1]^{\intercal}, & 451\leq k\leq 600.\end{cases}$$

Thus, the trajectory can be determined to be a square once the starting points are fixed.

The 2D field satisfies the following PDE

$$\frac{\partial z}{\partial t}=\theta(t)z(r,t), \tag{41}$$

where $\theta(t)=\frac{\cos(t)-e^{-t}}{\sin(t)+e^{-t}+2}$. Assume we have no knowledge of the coefficient $\theta(t)$ but we can obtain some history estimation to train the LSTM network. Assume we have 1000 sequential data points of $\theta(t)$, and the constant L in (31) is 10. Then we have 990 input-output pairs that can be split into 800 training data pairs and 190 testing data pairs. The network has one single LSTM layer of 200 units and is trained with learning rate of 0.005. After the LSTM network has been trained, it can be applied to (33) to estimate parameter $\theta(t)$ based on state estimation from the constrained cooperative Kalman filter (25)-(30). The constrained cooperative Kalman filter will update the state estimation based on estimated θ and new measurements.

Fig.3 shows the evolution of state estimation at the formation center along trajectory. Compared with the true values (red dashed lines), the state estimations (blue solid lines) have the same trend as the true values. The RMSE for field value z_c is 0.0770, and the RMSE for gradient estimation ∇z_c is $[0.0356,\ 0.0260]^{\intercal}$. This means that the proposed algorithm is capable of providing accurate state estimation.

The comparison between estimated parameter $\hat{\theta}$ and true value θ is shown in Fig.4. The RMSE for parameter estimation is 0.1577. We can observe that the estimated parameter follows the same trend as the true value, and relative large estimation error usually appears around the local maximum or minimum. By comparing Fig.3 with Fig.4, we can find that the time instance when parameter becomes maximum or minimum is very close to the time instance when the state reaches its maximum or minimum value. The state estimation error usually increases around such time instances, which leads to increased parameter estimation error.

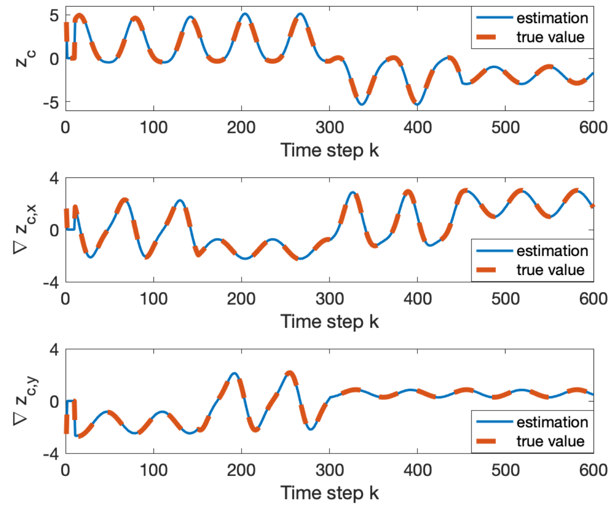

Fig. 3. State estimation of field value z_c, gradient $\nabla z_c=[\nabla z_{c,x},\nabla z_{c,y}]^{\intercal}$ at the formation center r_c along trajectory.

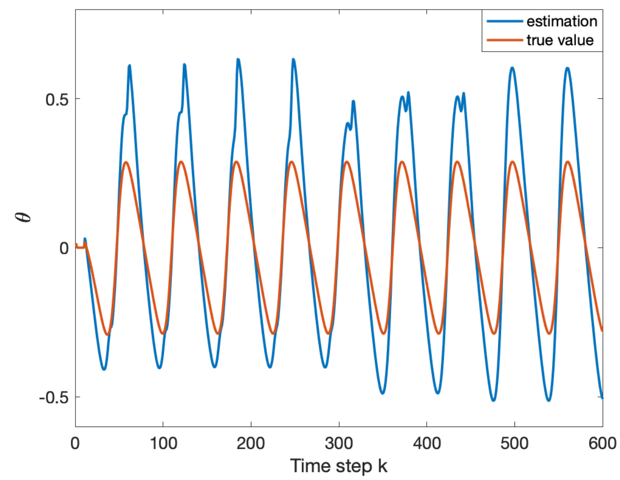

Fig. 4. Parameter estimation along trajectory.

VIII. Conclusions and Future Work

In this paper, we propose a constrained cooperative Kalman filter for state estimation using mobile sensor networks within a polynomial PDE field with provable convergence. Meanwhile, an LSTM-based Kalman filter has been applied to provide parameter estimation every time a new state estimation becomes available. In the future, we will combine the proposed algorithm with various control laws so that the mobile sensor networks can do source seeking, level curve tracking, predator avoidance, etc.

References

[1] A. Sommerfeld, *Partial differential equations in physics*. Academic press, 1949.

[2] H. Yu, Q. Gan, A. Bayen, and M. Krstic, "Pde traffic observer validated on freeway data," *IEEE Transactions on Control Systems Technology*, vol. 29, no. 3, pp. 1048–1060, 2021.

[3] S. Koga and M. Krstic, "Arctic sea ice state estimation from thermodynamic pde model," *Automatica*, vol. 112, p. 108713, 2020.

[4] J.-W. Wang, Y. Guo, M. Fahad, and B. Bingham, "Dynamic plume tracking by cooperative robots," *IEEE/ASME Transactions on Mechatronics*, vol. 24, no. 2, pp. 609–620, 2019.

[5] M. A. Demetriou, "Employing mobile sensor density to approximate state feedback kernels in static output feedback control of pdes," in *2021 American Control Conference (ACC)*. IEEE, 2021, pp. 2775–2781.

[6] H. Huang, A. V. Savkin, M. Ding, and C. Huang, "Mobile robots in wireless sensor networks: A survey on tasks," *Computer Networks*, vol. 148, pp. 1–19, 2019.

[7] P. K. Mishra, D. Chatterjee, and D. E. Quevedo, "Stochastic predictive control under intermittent observations and unreliable actions," *Automatica*, vol. 118, p. 109012, 2020.

[8] J. You, Y. Zhang, M. Li, K. Su, F. Zhang, and W. Wu, "Cooperative parameter identification of advection-diffusion processes using a mobile sensor network," in *2017 American Control Conference (ACC)*, 2017, pp. 3230–3236.

[9] J. You and W. Wu, "Online passive identifier for spatially distributed systems using mobile sensor networks," *IEEE Transactions on Control Systems Technology*, vol. 25, no. 6, pp. 2151–2159, 2017.

[10] W. Wu, J. You, Y. Zhang, M. Li, and K. Su, "Parameter identification of spatial–temporal varying processes by a multi-robot system in realistic diffusion fields," *Robotica*, vol. 39, no. 5, p. 842–861, 2021.

[11] M. Krstic and A. Smyshlyaev, "Adaptive control of pdes," *Annual Reviews in Control*, vol. 32, no. 2, pp. 149–160, 2008.

[12] S. Khatibi, G. O. Cassol, and S. Dubljevic, "Linear model predictive control for a cascade ode-pde system," in *2020 American Control Conference (ACC)*. IEEE, 2020, pp. 4521–4526.

[13] Y. Wen, X. Lou, W. Wu, and B. Cui, "Impulsive adaptive observer design for a class of hybrid ode–pde cascade systems with uncertain parameters," *Systems & Control Letters*, vol. 154, p. 104969, 2021.

[14] I. Karafyllis, M. Krstic, and K. Chrysafi, "Adaptive boundary control of constant-parameter reaction–diffusion pdes using regulation-triggered finite-time identification," *Automatica*, vol. 103, pp. 166–179, 2019.

[15] X. Xun, J. Cao, B. Mallick, A. Maity, and R. J. Carroll, "Parameter estimation of partial differential equation models," *Journal of the American Statistical Association*, vol. 108, no. 503, pp. 1009–1020, 2013.

[26] S. Hochreiter and J. Schmidhuber, "Long short-term memory," *Neural Computation*, vol. 9, no. 8, pp. 1735–1780, 1997.

[16] J. You, F. Zhang, and W. Wu, "Cooperative filtering for parameter identification of diffusion processes," in *2016 IEEE 55th Conference on Decision and Control (CDC)*, 2016, pp. 4327–4333.

[17] Z. Zhang, M. Hou, F. Zhang, and C. R. Edwards, "An lstm based kalman filter for spatio-temporal ocean currents assimilation," in *Proceedings of the International Conference on Underwater Networks & Systems*, 2019, pp. 1–7.

[18] D. Simon and T. L. Chia, "Kalman filtering with state equality constraints," *IEEE transactions on Aerospace and Electronic Systems*, vol. 38, no. 1, pp. 128–136, 2002.

[19] D. Simon, "Kalman filtering with state constraints: a survey of linear and nonlinear algorithms," *IET Control Theory & Applications*, vol. 4, no. 8, pp. 1303–1318, 2010.

[20] F. Zhang and N. E. Leonard, "Cooperative filters and control for cooperative exploration," *IEEE Transactions on Automatic Control*, vol. 55, no. 3, pp. 650–663, 2010.

[21] R. Drummond and Y. Zheng, "Impact of disturbances on mixed traffic control with autonomous vehicles," in *2020 59th IEEE Conference on Decision and Control (CDC)*, 2020, pp. 220–225.

[22] B. T. Agyeman, S. Bo, S. R. Sahoo, X. Yin, J. Liu, and S. L. Shah, "Soil moisture map construction by sequential data assimilation using an extended kalman filter," in *2021 American Control Conference (ACC)*, 2021, pp. 4351–4356.

[23] M. Lin, C. Yang, and D. Li, "Hybrid strategy based model parameter estimation of irregular-shaped underwater vehicles for predicting velocity," *Robotics and Autonomous Systems*, vol. 127, p. 103480, 2020.

[24] P. Skarin, J. Eker, and K.-E. Årzén, "A cloud-enabled rate-switching mpc architecture," in *2020 59th IEEE Conference on Decision and Control (CDC)*, 2020, pp. 3151–3158.

[25] P. Skarin, J. Eker, and K.-E. Årzén, "Cloud-based model predictive control with variable horizon," *IFAC-PapersOnLine*, vol. 53, no. 2, pp. 6993–7000, 2020.

[27] A. Sherstinsky, "Fundamentals of recurrent neural network (rnn) and long short-term memory (lstm) network," *Physica D: Nonlinear Phenomena*, vol. 404, p. 132306, 2020.

[28] A. H. Jazwinski, *Stochastic Processes and Filtering Theory*. Academic Press, 1970.