



# Distributed Nonlinear Kalman Filter with Communication Protocol

## Document Version

Accepted author manuscript

[Link to publication record in Manchester Research Explorer](#)

## Citation for published version (APA):

Tnunay, H. I., Li, Z., & Ding, Z. (2019). Distributed Nonlinear Kalman Filter with Communication Protocol. *Information Sciences*, 513, 270-288.  
<https://www.sciencedirect.com/science/article/pii/S0020025519310229?via%3Dihub>

## Published in:

Information Sciences

## Citing this paper

Please note that where the full-text provided on Manchester Research Explorer is the Author Accepted Manuscript or Proof version this may differ from the final Published version. If citing, it is advised that you check and use the publisher's definitive version.

## General rights

Copyright and moral rights for the publications made accessible in the Research Explorer are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

## Takedown policy

If you believe that this document breaches copyright please refer to the University of Manchester's Takedown Procedures [<http://man.ac.uk/04Y6Bo>] or contact [uml.scholarlycommunications@manchester.ac.uk](mailto:uml.scholarlycommunications@manchester.ac.uk) providing relevant details, so we can investigate your claim.



# Distributed Nonlinear Kalman Filter with Communication Protocol\*

Hilton Tnunay<sup>a</sup>, Zhenhong Li<sup>b</sup>, Zhengtao Ding<sup>a,\*</sup>

<sup>a</sup>*Department of Electrical and Electronic Engineering, The University of Manchester,  
Manchester M13 9PL, UK*

<sup>b</sup>*School of Electronic and Electrical Engineering, University of Leeds, Leeds LS2 9JT, UK*

---

## Abstract

This paper proposes an optimal design of the general distributed nonlinear Kalman-based filtering algorithm to tackle the discrete-time estimation problem with noisy communication networks. The algorithm extends the Kalman filter by enabling it to predict the noisy communication data and fuse it with the received neighboring information to produce a posterior estimate value. In the prediction step, the unscented transformations of the estimate values and covariances originated in the Unscented Kalman Filter (UKF) are exploited. In the update step, a communication protocol is appended to the posterior estimator, which consequently leads to a modified posterior error covariance containing the covariance of the communication term with its communication gain. Both Kalman and communication gains are then optimised to collectively minimise the mean-squared estimation error. Afterwards, stochastic stability analysis is performed to guarantee its exponential boundedness. To exemplify the performance, this algorithm is applied to a group of robots in a sensor network assigned to estimate an unknown information distribution over an area in the optimal coverage control problem. Comparative numerical experiments finally verify the effectiveness of our design.

**Keywords:** Distributed nonlinear Kalman filter, nonlinear estimation, communication protocol, sensor network, optimal coverage control.

---

\*This work is supported by Indonesia Endowment Fund for Education (LPDP), Ministry of Finance, Indonesia; and the Science and Technology Facilities Council under Grant ST/N006852/1.

\*Corresponding author

*Email addresses:* `hilton.tnunay@manchester.ac.uk` (Hilton Tnunay),  
`z.h.li@leeds.ac.uk` (Zhenhong Li), `zhengtao.ding@manchester.ac.uk`  
(Zhengtao Ding)

---

## 1. Introduction

The field of Kalman filters in a connected sensor network has become an interesting area to investigate. It incorporates the concept of distributed control systems and the Kalman filtering method. In the case of centralised Kalman filters, for both the linear and nonlinear systems, intensive investigations have been carried out, for example, by [1], [2], [3], [4], [5], [6] and [7]. While in the case of networked linear filter, the study reported by [8] added updating weight on the measurement dynamics to fuse information from a number of sensors, while [9] includes the low, high and band-pass consensus filters. The stability of a networked filter, namely, the Kalman-Consensus filter, has been analysed in [10], while various scenarios of diffusive Kalman filtering for the linear system have been designed in [11]. Generalising the concept of distributed Kalman filtering, the networked nonlinear Kalman filtering has also attracted numerous developments in distributed Extended Kalman Filter (EKF), Unscented Kalman Filter (UKF) and Cubature Kalman Filter (CKF) [12], [13], [14] and [15].

Compared to the centralised Kalman filter, the distributed version has attracted more attentions due to the high scalability, robustness to failure, and flexibility [15], [16]. The distributed protocol is able to reduce the computation burden of a central processor while maintaining the performance of the global result. In practice, communication medium has noises which can affect the quality of shared information among agents. Although there have been some works on the distributed nonlinear Kalman filter algorithms in literature, such as in [9], [15], the communication noise has not been considered and analysed. Both in the linear and nonlinear cases, the previous works on distributed Kalman filter have considered the implementation of consensus protocol to minimise the disagreement of the estimate results among the sensors. However, the optimal gain of the consensus term has not been analysed in order to guarantee the performance. If one fails to pick the consensus gain, the result may oscillate and become unstable. Thus, the optimal Kalman and consensus gains need to be designed such that the optimality and boundedness can be guaranteed.

In the application to mobile sensor network, sensors can be referred to as a group of robots, each carrying a sensor, assigned to measure a physical quantity over an environment – which, for example, can be applied to find the source of radiation. One of the important applications in the sensor network is the optimal sensor coverage – which also leads to the optimised sensor configuration. Loca-

tional optimisation, originated from the field of operation research aims to find the best locations of agents given an interest function [17], [18]. Centroidal Voronoi Tessellation has become a recognised tool to solve this locational problem [19], [20]. Cooperative control design is important to drive robots towards the optimal locations. The related control protocol has been reported in [20], [21] and [22], which iteratively minimises the objective function of the locational optimisation. The information of a covered area is represented by a density function which in reality is unknown but can be estimated by the agents, this is solved in field estimation problem. Studies about coverage problems with field estimation of unknown information distribution can be found in [23], [24], and [25]. These field estimation algorithms have not considered the noises in the measurement and communication network. Moreover, these algorithms require some gains of the estimator to be tuned prior to execution and the upper boundary of the gains has not been given. Failure to choose the suitable gains prior to the execution might lead to an unstable system.

In this paper, we present an optimal design of nonlinear Kalman filter to estimate the state of a dynamical process in a distributed manner using the shared information among the agents. Different from existing techniques in literature, such as in [9], [12], [13], [14] and [15], our algorithm generalises the distributed Kalman filter to accomodate any communication mechanism which uses not only the measurement from an agent's own sensor as consideration for estimating the process, but also the shared information from the neighboring agents. In addition to standard Kalman filter, this distributed filtering algorithm also considers noisy communicated information within the network. The proposed technique alters the procedure in the prediction and update steps of the unscented Kalman filter. In the prediction, the prior state, measurement and communication estimates and their covariances are designed utilising the sigma points calculated from the previous estimate and covariance. While, in the subsequent update step, the posterior estimator of this filter comprises the prior estimate term, measurement correction term and communication correction term - leading to the posterior covariance containing the prior state, measurement and communication covariances. Afterwards, the Kalman and communication gains of the estimator are optimally designed. It can be said that solving the optimal gain of existing distributed Kalman filter with consensus algorithm is a special case of this technique. To verify the effectiveness of the proposed algorithm, it is then applied to estimate the information distribution of optimal coverage problem in an area by a sensor network. Different with the state-of-the-art field estimation algorithm reported in [25], the proposed filter in this paper has considered the noises in measurement and in the communica-

tion network and also optimised the estimator gains in every iteration to avoid instability of the system caused by failure to choose the appropriate gains.

The remainder of this paper is structured as follows. In Section II, some necessary mathematical notations and graph theory are reviewed. Section III declares the formulation of the general distributed estimation problem. Section IV presents the result on the distributed Kalman-based nonlinear estimation, followed by Section V discussing the application of the proposed filter to the field estimation of the optimal coverage problem. Afterwards, in Section VI, comparative simulations on optimal coverage problem using distributed Kalman filter algorithm and the modified-consensus observer in [25] are presented to validate the performance of the proposed methods. Finally, Section VII concludes this work.

## 2. Preliminaries

### 2.1. Notations

Throughout this paper, the following notations are used. We use  $\mathbb{R}$ ,  $\mathbb{R}^+$ ,  $\mathbb{R}^n$  and  $\mathbb{R}^{n \times m}$ , respectively, to denote the set of real numbers, positive real numbers,  $n$ -dimension vectors with real-valued entries, and  $n \times m$ -matrices with real-valued entries. Related to integer-valued sets,  $\mathbb{Z}$ ,  $\mathbb{Z}^+$  and  $\mathbb{Z}^*$  are to denote the set of integers, positive integers and non-negative integers, respectively.

For matrix and vector,  $I_n$ ,  $0_n$  and  $1_n$  denote the  $n \times n$  identity matrix,  $n$ -sized column vector of zero and one, respectively. Let  $a \in \mathbb{R}^n$  and  $A \in \mathbb{R}^{n \times m}$ . Then,  $\|a\|$  is the Euclidean norm of  $a$ ;  $\text{tr}(A)$  is the trace of matrix  $A$ ; and  $\text{diag}(a)$  is the diagonal matrix constructed by the elements of  $a$ . Let  $a = \{a_1, a_2, \dots, a_p\}$ , for  $p > 0$  and  $a_q \in \mathbb{R}^r$ , then  $\text{vec}_p(a_q) \in \mathbb{R}^{pr}$  denotes a vector combining the elements of  $a$ , with  $1 \leq q \leq p$ . In this work, we use operator  $\otimes$  for the Kronecker product of two matrices.

Related to statistics, define  $X \in \mathbb{R}^{n \times m}$ , for  $n, m \in \mathbb{Z}^+$ , as a random variable. The statistical notation for expected value of  $X$  is given by  $\mathbb{E}[X]$  and the covariance of  $X$  by  $\text{cov}(X)$ . Correspondingly, the notation to express Gaussian distribution is in the form of  $\mathcal{N}(\mu, \Sigma)$ , where  $\mu = \mathbb{E}[X]$  and  $\Sigma = \text{cov}(X)$ .

Since we work on the discrete-time domain, some discrete-time, index-free notations are used for convenience. Let  $\vartheta_k \in \mathbb{R}^d$  be any discrete-time based variable at time step  $k$ , for  $k \in \mathbb{Z}^*$ . Then, to suppress the expressions, we have  $\vartheta$  as the substitute of  $\vartheta(k)$ ,  $\vartheta^+$  of  $\vartheta(k+1)$ , and  $\vartheta^-$  of  $\vartheta(k-1)$ .

## 2.2. Graph Theory

We now review some useful notations and definitions of graph theory for a network and control topology. A graph  $\mathcal{G}(\mathcal{V}, \mathcal{E})$  is a collection of vertices  $\mathcal{V}$  connected by a collection of edges  $\mathcal{E} \subseteq \mathcal{V} \times \mathcal{V}$ . If there exists an edge  $(i, j) \in \mathcal{E}$ , agent  $i$  is able to receive information from agent  $j$ . If, for  $(i, j) \in \mathcal{E}$ , there exists  $(j, i) \in \mathcal{E}$ , the graph is called undirected. We refer  $j \in N_i \subset \mathcal{V}$ , for  $j \neq i$ , to the neighbor of agent  $i$  if  $(i, j) \in \mathcal{E}$ .

To algebraically express the connectivity of a graph with  $n$  vertices, we entail the concept of adjacency, in-degree and Laplacian matrices. The adjacency matrix, denoted as  $\mathcal{A} = [a_{ij}] \in \mathbb{R}^{n \times n}$ , is a square matrix whose elements are given by  $a_{ij} = 1$  if  $(j, i) \in \mathcal{E}$ , and  $a_{ij} = 0$ , otherwise. If a graph is undirected, it is straightforward to conclude that  $\mathcal{A}$  is a symmetric matrix. Correspondingly, summing up the  $i$ -th row of  $\mathcal{A}$  yields an in-degree matrix, i.e.,  $\mathcal{D} = \text{diag}(d_i) \in \mathbb{R}^{n \times n}$  such that  $d_i = \sum_{j \in N_i} a_{ij}$ , for all  $i \in V$ . The associated Laplacian is subsequently defined as  $\mathcal{L} = [\mathcal{L}_{ij}] \in \mathbb{R}^{n \times n}$ , for  $\mathcal{L}_{ii} = d_i$  and  $\mathcal{L}_{ij} = -a_{ij}$ . Using arithmetical operation, a Laplacian matrix can be obtained using  $\mathcal{L} = \mathcal{D} - \mathcal{A}$ . The eigenvalues of Laplacian matrix  $\mathcal{L}$  can sequentially be expressed as  $0 = \lambda_1 \leq \lambda_2 \leq \lambda_3 \leq \dots \leq \lambda_n$ . For a strongly-connected directed graph, there is only one zero eigenvalue of  $\mathcal{L}$ , i.e.,  $\lambda_1 = 0$ . Moreover, according to Courant-Fisher theorem, the smallest and largest non-zero eigenvalues of  $\mathcal{L}$  satisfy  $\lambda_2 \|z\|^2 \leq z^T \mathcal{L} z$  for  $1_n^T z = 0$ , and  $\lambda_n \|z\|^2 \geq z^T \mathcal{L} z$  for  $z \in \mathbb{R}^n$ , respectively [26].

## 3. Problem Formulation

Consider a group of  $n$  agents connected in a network whose topology is represented by a graph  $\mathcal{G}_n(\mathcal{V}_n, \mathcal{E}_n)$ , for  $\mathcal{V}_n = \{1, 2, \dots, n\}$  and  $\mathcal{E}_n \subseteq \mathcal{V}_n \times \mathcal{V}_n$ . These agents are assigned to estimate a dynamical process represented by a function  $f : \mathbb{R}^p \times \mathbb{R}^p \rightarrow \mathbb{R}^p$ , such that

$$\xi^+ = f(\xi, \xi_v), \quad (1)$$

where  $\xi \in \mathbb{R}^p$  and  $v \in \mathbb{R}^p$ , respectively, denote the unknown state, and the process noise at time step  $k$ , with  $k \in \mathbb{Z}^*$ .

To estimate the process (1) using a group of networked agents in a distributed manner, each agent is equipped with a sensor whose measurement data is modelled by a function  $h_i : \mathbb{R}^p \times \mathbb{R}^q \rightarrow \mathbb{R}^q$  such that

$$\zeta_i = h_i(\xi, w_i), \text{ for } i \in \mathcal{V}_n, \quad (2)$$

where  $\zeta_i \in \mathbb{R}^q$  and  $w_i \in \mathbb{R}^q$ , respectively, denote the sensing data and the measurement noise at time step  $k$ .

Since each agent within a network requires information from its neighbours to collaboratively estimate a dynamical process, we need to model the communication network of the agents and define how each agent use the shared information. As a proposed mechanism in the distributed nonlinear Kalman filtering, the communication model of a noisy network at time step  $k$  is represented by a function  $G : \mathbb{R}^{np} \times \mathbb{R}^{nr} \rightarrow \mathbb{R}^{nr}$ , i.e.,

$$z = G(x, s), \quad (3)$$

where, in this communication model,  $x = 1_n \otimes \xi$ ,  $z = \text{vec}_n(z_i) \in \mathbb{R}^{nr}$  and  $s = \text{vec}_n(s_i) \in \mathbb{R}^{nr}$  are the augmented state to estimate, the communicated data and the communication noise, respectively.

**Remark 1.** *The communication model of a network provides a representation of the connection topology of a network. For example, consider a group of  $n$  robots are assigned to estimate  $p$  unknown states,  $\xi \in \mathbb{R}^p$ , in which the data received by agent  $i$  is denoted by  $z_i \in \mathbb{R}^p$  and the communication noise by  $s_i \in \mathbb{R}^r$ . Suppose that the communication topology is modelled via a graph with Laplacian matrix  $\mathcal{L} \in \mathbb{R}^{np \times np}$ . Then, using the defined model, the communication process can be expressed as  $z = \mathcal{L}x + s$ .*

Throughout this paper, to improve the brevity of the filter design and formulation, augmented expressions of the process, measurement and communication dynamics are equivalently described as

$$x^+ = F(x, v) \quad (4)$$

$$y = H(x, w) \quad (5)$$

$$z = G(x, s) \quad (6)$$

where  $y = \text{vec}_n(\zeta_i)$ ,  $F = \text{vec}_n(f_i)$ ,  $H = \text{vec}_n(h_i)$ ,  $v = 1_n \otimes \xi_v$ , and  $w = \text{vec}_n(w_i)$ .

There are some assumptions made regarding the noises, given as follows.

**Assumption 1.** *The expected values of process, measurement and communication noises are zero, i.e.,  $E[v] = 0$ ,  $E[w] = 0$  and  $E[s] = 0$ .*

**Assumption 2.** *The covariances of process, measurement and communication noises are  $E[vv^\top] = Q$ ,  $E[ww^\top] = R$ , and  $E[ss^\top] = S$ , where  $Q_{jk} = q_{jk}\delta_{jk}$ ,  $R_{jk} = r_{jk}\delta_{jk}$ , and  $S_{jk} = s_{jk}\delta_{jk}$ , for some scalars  $q_{jk}, r_{jk}, s_{jk} \in \mathbb{R}$  and the Kronecker delta  $\delta_{jk}$ .*

**Assumption 3.** *The process, measurement and communication noises are uncorrelated, i.e.,  $E[vw^\top] = 0$ ,  $E[vs^\top] = 0$ ,  $E[ws^\top] = 0$ .*

**Assumption 4.** *The process, measurement and communication estimates of agent  $i$  and  $j$ , for  $i \neq j$ , are uncorrelated, i.e., their cross-covariance matrices are zero.*

The objective of the distributed estimation is to minimise the collective mean-squared error (MSE) of the estimation of the agents. The performance index is formulated as

$$\min_{\hat{x}} E [(x - \hat{x})^\top (x - \hat{x})] \quad (7)$$

where  $\hat{x} = \text{vec}_n(\hat{x}_i)$  is the estimate value of  $x$ . In the following filter design, the estimate value of  $x$  refers to the posterior estimate of the Kalman filter.

To show the boundedness of the filtering algorithms, we need the lemma about the stability of a stochastic process which was given in literature, such as [27], [12].

**Lemma 1** (Stochastic Boundedness). *If, for  $\theta \in \mathbb{R}^d$  being a stochastic process at time step  $k$ ,  $k \in \mathbb{Z}^*$ , there exists a scalar stochastic process  $V(\varepsilon)$  with  $\varepsilon = \theta - \theta'$  satisfying these conditions:*

1.  $\underline{v}\|\varepsilon\|^2 \leq V(\varepsilon) \leq \bar{v}\|\varepsilon\|^2$ , for  $\underline{v}, \bar{v} > 0$ ,
2.  $E[V^+(\varepsilon^+)|\varepsilon] \leq \mu + (1 - \sigma)V(\varepsilon)$ , for  $\mu > 0$ ,  $0 \leq \sigma \leq 1$ .

*Then, the stochastic process  $\varepsilon$  is exponentially bounded in mean square such that*

$$E [\|\varepsilon\|^2] \leq \frac{\bar{v}}{\underline{v}} E[\|\varepsilon_0\|^2] (1 - \sigma)^k + \frac{\mu}{\underline{v}} \sum_{l=1}^{k-1} (1 - \sigma)^l. \quad (8)$$

There are some useful variables that are necessary for the filter design.  $\hat{x}$ ,  $\bar{x}$ ,  $\bar{y}$ , and  $\bar{z}$ , respectively, denote the posterior state estimate of  $x$ , prior state estimate of  $x$ , prior measurement estimate of  $y$ , and prior communication estimate of  $z$  at time step  $k$ , for  $k \in \mathbb{Z}^*$ . Related to the covariance update,  $P$  denotes the posterior covariance of the error of the posterior estimates;  $\bar{P}_{xx}$ ,  $\bar{P}_{yy}$ ,  $\bar{P}_{zz}$ ,  $\bar{P}_{xy}$ ,  $\bar{P}_{xz}$ , and  $\bar{P}_{yz}$ , respectively, denote the prior state-to-state, measurement-to-measurement, communication-to-communication, state-to-measurement, state-to-communication and measurement-to-communication covariances.



## 4. Distributed Nonlinear Filter Design

In this filter design, the state, measurement and communication models are given in (4), (5) and (6), respectively, and these functions are considered to be nonlinear.

### 4.1. Prediction

The means of unscented transformation used in the Unscented Kalman Filter is modified in the prediction step to produce the predicted state, measurement and communication estimates and their covariances.

Define  $\tilde{x} = [\hat{x}^\top v^\top w^\top s^\top]^\top$  as a vector of random variables augmenting the state and noises; while  $\tilde{P} = \text{diag}(P, Q, R, S)$  as the augmented covariance at time step  $k$ . The augmented form of sigma vectors at time step  $k - 1$  is defined as  $\mathcal{X}^- = [(\mathcal{X}^{x-})^\top (\mathcal{X}^{v-})^\top (\mathcal{X}^{w-})^\top (\mathcal{X}^{s-})^\top]^\top \in \mathbb{R}^{a \times (2a+1)}$ ,  $a = (2p + q + r)n$ . Its entries are given by

$$\mathcal{X}^- = \begin{bmatrix} \tilde{x}^- & (1_a^\top \otimes \tilde{x}^-) + \sqrt{\kappa \tilde{P}^-} & (1_a^\top \otimes \tilde{x}^-) - \sqrt{\kappa \tilde{P}^-} \end{bmatrix}, \quad (9)$$

with  $\kappa = a + b$ . Expression  $1_a^\top \otimes \tilde{x}^-$  augments the previous posterior estimates to comply with the dimension of the augmented state covariance. Here,  $b$  is a scaling parameter expressed as  $b = \varrho_1^2(a + \varrho_2) - a$ , where  $\varrho_1$  correlates with the spread of sigma points around  $\tilde{x}^-$ , while  $\varrho_2$  is another scaling parameter. The details regarding the criterion of these parameters are provided in [28].

The sigma vector is subsequently transformed through the nonlinear functions. For every  $l$ -th column of  $\mathcal{X}^-$ , with  $l \in \{0, 1, 2, \dots, 2a\}$ , the sigma vector is mapped through (4) such that the prior sigma vector of the state can be written as

$$\bar{\mathcal{X}}_l = F(\mathcal{X}_l^{x-}, \mathcal{X}_l^{v-}). \quad (10)$$

The prior estimate  $\bar{x}$  and covariance  $\bar{P}_{xx}$  are attained using the approximated weighted mean and covariance of the sigma points in the form of

$$\bar{x} = \sum_{l=0}^{2a} W_l^m \bar{\mathcal{X}}_l, \quad (11)$$

$$\bar{P}_{xx} = \sum_{l=0}^{2a} W_l^c [\bar{\mathcal{X}}_l - \bar{x}] [\bar{\mathcal{X}}_l - \bar{x}]^\top, \quad (12)$$

where the weights are given by

$$\begin{aligned} W_0^m &= b/\kappa, \\ W_0^c &= b/\kappa + (1 - \varrho_1^2 + \varrho_2), \\ W_l^m &= W_l^c = 1/(2\kappa), \quad 0 < l \leq 2a. \end{aligned}$$

To predict the measurement data, the sigma vector of the state in (10) is also mapped through the nonlinear measurement function (5) such that the sigma vector of the measurement is formulated as

$$\bar{\mathcal{Y}}_l = H(\bar{\mathcal{X}}_l, \mathcal{X}_l^{w-}). \quad (14)$$

Accordingly, the prior measurement estimate and its covariance,  $\bar{y}$  and  $\bar{P}_{yy}$ , are calculated using the approximation of weighted mean and covariance, i.e.,

$$\bar{y} = \sum_{l=0}^{2a} W_l^m \bar{\mathcal{Y}}_l, \quad (15)$$

$$\bar{P}_{yy} = \sum_{l=0}^{2a} W_l^c [\bar{\mathcal{Y}}_l - \bar{y}] [\bar{\mathcal{Y}}_l - \bar{y}]^\top, \quad (16)$$

respectively.

The prediction of the communication data and its covariance can also be attained by mapping the sigma vector (10) to the communication function (6), yielding the communication sigma vector given by

$$\bar{\mathcal{Z}}_l = G(\bar{\mathcal{X}}_l, \mathcal{X}_l^{s-}). \quad (17)$$

The associated approximation of the weighted mean and covariance can be expressed as

$$\bar{z} = \sum_{l=0}^{2a} W_l^m \bar{\mathcal{Z}}_l, \quad (18)$$

$$\bar{P}_{zz} = \sum_{l=0}^{2a} W_l^c [\bar{\mathcal{Z}}_l - \bar{z}] [\bar{\mathcal{Z}}_l - \bar{z}]^\top. \quad (19)$$

The prior cross covariance matrices between the state estimate and measurement data, between state estimate and communication data, and between measurement and communication data can, respectively, also be formulated utilising the

sigma vectors declared previously in the form of

$$\bar{P}_{xy} = \sum_{l=0}^{2a} W_l^c [\bar{\mathcal{X}}_l - \bar{x}] [\bar{\mathcal{Y}}_l - \bar{y}]^\top, \quad (20)$$

$$\bar{P}_{xz} = \sum_{l=0}^{2a} W_l^c [\bar{\mathcal{X}}_l - \bar{x}] [\bar{\mathcal{Z}}_l - \bar{z}]^\top, \quad (21)$$

and

$$\bar{P}_{yz} = \sum_{l=0}^{2a} W_l^c [\bar{\mathcal{Y}}_l - \bar{y}] [\bar{\mathcal{Z}}_l - \bar{z}]^\top. \quad (22)$$

**Remark 2.** Notice that the prior state, measurement and communication covariances are symmetric matrices, i.e.,  $\bar{P}_{xx} = \bar{P}_{xx}^\top$ ,  $\bar{P}_{yy} = \bar{P}_{yy}^\top$ , and  $\bar{P}_{zz} = \bar{P}_{zz}^\top$ .

#### 4.2. Update

In this paper, an estimator at time step  $k$  append the additional communication term as a correcting parameter to ensure the collective estimate are achieved. The proposed estimator is formulated as

$$\hat{x} = \bar{x} + K(y - \bar{y}) + C(z - \bar{z}). \quad (23)$$

This estimator comprises three terms: the first is the prior state estimate; the second term is the conditional-mean update of Bayesian estimation method to interfere the measurement data with the prior measurement estimate; and the last term is the proposed communication term to correct the predicted communication data with the information obtained from neighboring agents. The parameters  $K = \text{diag}_n(K_i) \in \mathbb{R}^{np \times nq}$  and  $C = \text{diag}_n(C_i) \in \mathbb{R}^{np \times nr}$ , respectively, denote the Kalman gain and communication gains. In order to minimise the cost function (7), one has to find the optimal value of these gains.

Recall the cost function in (7) which can equivalently be expressed as

$$\mathbb{E} [(x - \hat{x})^\top (x - \hat{x})] = \mathbb{E} [\text{tr} ((x - \hat{x})(x - \hat{x})^\top)], \quad (24)$$

in which  $\mathbb{E}[\text{tr}\{u\}] = \text{tr}\{\mathbb{E}[u]\}$  and  $\text{tr}\{u^\top u\} = \text{tr}\{uu^\top\}$  have been applied. Based on (24), minimising the mean-squared error is equivalent to minimising the trace of the covariance of the estimate error, i.e.,

$$P = \mathbb{E} [(x - \hat{x})(x - \hat{x})^\top]. \quad (25)$$

Let  $\bar{e}_x = x - \bar{x}$  denote the error between the unknown state and the prior estimate;  $\bar{e}_y = y - \bar{y}$  the error between the measurement data and its prior estimate; and  $\bar{e}_z = z - \bar{z}$  the error between communication data and its prior estimate. Using these error parameters, the error of the posterior estimate  $\hat{e}_x = x - \hat{x}$  is given by

$$\hat{e}_x = \bar{e}_x - K\bar{e}_y - C\bar{e}_z \quad (26)$$

where the posterior estimate in (23) has been inserted. Afterwards, substituting (26) to the posterior covariance in (25) and applying the linearity property of expected value result in the posterior covariance of estimate error:

$$\begin{aligned} P = & \mathbb{E} [\bar{e}_x \bar{e}_x^\top] - K \mathbb{E} [\bar{e}_y \bar{e}_x^\top] - \mathbb{E} [\bar{e}_x \bar{e}_y^\top] K^\top - C \mathbb{E} [\bar{e}_z \bar{e}_x^\top] - \mathbb{E} [\bar{e}_x \bar{e}_z^\top] C^\top \\ & + K \mathbb{E} [\bar{e}_x \bar{e}_y^\top] K^\top + C \mathbb{E} [\bar{e}_y \bar{e}_z^\top] C^\top + K \mathbb{E} [\bar{e}_y \bar{e}_z^\top] C^\top + C \mathbb{E} [\bar{e}_z \bar{e}_y^\top] K^\top. \end{aligned} \quad (27)$$

Notice that  $\bar{P}_{xx} = \mathbb{E}[\bar{e}_x \bar{e}_x^\top]$ ,  $\bar{P}_{yy} = \mathbb{E}[\bar{e}_y \bar{e}_y^\top]$ ,  $\bar{P}_{zz} = \mathbb{E}[\bar{e}_z \bar{e}_z^\top]$ ,  $\bar{P}_{xy} = \mathbb{E}[\bar{e}_x \bar{e}_y^\top]$ ,  $\bar{P}_{xz} = \mathbb{E}[\bar{e}_x \bar{e}_z^\top]$ ,  $\bar{P}_{yz} = \mathbb{E}[\bar{e}_y \bar{e}_z^\top]$ . Also, notice that  $\bar{P}_{xy}^\top = \bar{P}_{yx}$ ,  $\bar{P}_{xz}^\top = \bar{P}_{zx}$ ,  $\bar{P}_{yz}^\top = \bar{P}_{zy}$ . Thus, it is straightforward to conclude that the covariance matrix can be formulated as

$$\begin{aligned} P = & \bar{P}_{xx} + K\bar{P}_{yy}K^\top + C\bar{P}_{zz}C^\top + K\bar{P}_{yz}C^\top + C\bar{P}_{yz}^\top K^\top \\ & - K\bar{P}_{xy}^\top - \bar{P}_{xy}K^\top - C\bar{P}_{xz}^\top - \bar{P}_{xz}C^\top \end{aligned} \quad (28)$$

The result about the gains update is stated in the following theorem.

**Theorem 1** (Kalman and Communication Gains Update). *Let (7) be the objective function of the distributed nonlinear Kalman filter. Let the posterior estimate be given by (23) and the covariance by (28). Then, the optimal communication and Kalman gains at time step  $k \in \mathbb{Z}^*$  are, respectively, formulated by*

$$C = \bar{P}_{xz} \bar{P}_{zz}^{-1}, \quad (29)$$

$$K = (\bar{P}_{xy} - C\bar{P}_{yz}^\top) \bar{P}_{yy}^{-1}. \quad (30)$$

Moreover, with these gains, the covariance in (28) can be expressed as

$$P = \bar{P}_{xx} - K\bar{P}_{yy}K^\top - C\bar{P}_{zz}C^\top. \quad (31)$$

*Proof.* Recall the cost function in (7). This cost function can also be written as

$$\mathbb{E} [(x - \hat{x})^\top (x - \hat{x})] = \mathbb{E} [\text{tr}\{(x - \hat{x})(x - \hat{x})^\top\}], \quad (32)$$

in which  $E[\text{tr}\{u\}] = \text{tr}\{E[u]\}$  and  $\text{tr}\{u^\top u\} = \text{tr}\{uu^\top\}$  have been applied. Since minimising the mean-squared error (32) is equivalent to minimising the trace of the covariance of the estimate error, the optimal Kalman and communication gain can be derived by taking the derivative of the trace of posterior covariance matrix with respect to the Kalman and communication gain. Thus, taking the derivative of (28) with respect to the Kalman gain gives

$$\frac{\partial \text{tr}(P)}{\partial K} = K\bar{P}_{yy} + C\bar{P}_{yz}^\top - \bar{P}_{xy} = 0$$

which leads to

$$\bar{P}_{xy} = K\bar{P}_{yy} + C\bar{P}_{yz}^\top. \quad (33)$$

Plugging (33) into (28) yields

$$P = \bar{P}_{xx} - K\bar{P}_{yy}K^\top + C\bar{P}_{zz}C^\top - C\bar{P}_{xz}^\top - \bar{P}_{xz}C^\top. \quad (34)$$

Thus, taking the derivative of (34) with respect to  $C$ , i.e.,

$$\frac{\partial \text{tr}(P)}{\partial C} = C\bar{P}_{zz} - \bar{P}_{xz} = 0$$

results in the Communication gain given by

$$C = \bar{P}_{xz}\bar{P}_{zz}^{-1}. \quad (35)$$

Moreover, rearranging (35) to obtain  $\bar{P}_{xz}$  and plugging it to (34) give the compact covariance expressed as

$$P = \bar{P}_{xx} - K\bar{P}_{yy}K^\top - C\bar{P}_{zz}C^\top. \quad (36)$$

This completes the proof.  $\square$

#### 4.3. Instrumental Matrices

To analyse the performance of the estimation algorithm we have designed, some additional matrices are utilised – which has also been used in Kalman-filter based performance analysis such as in [29]. In the following analysis,  $\bar{e}_x = x - \bar{x}$ ,  $\hat{e}_x = x - \hat{x}$ ,  $\bar{e}_y = y - \bar{y}$ , and  $\bar{e}_z = z - \bar{z}$  denote as the prior, posterior estimate, measurement and communication errors at time step  $k \in \mathbb{Z}^*$ , respectively.

Since  $x$  depends on the state estimate at  $k-1$ , by employing Taylor expansion to  $x$  in (4) around  $\hat{x}^-$ , we have

$$x = F(\hat{x}^-) + \nabla F(\hat{x}^-)\hat{e}_x^- + \frac{1}{2}\nabla^2 F(\hat{x}^-)(\hat{e}_x^-)^2 + \cdots + v^-, \quad (37)$$

where

$$\nabla^n F(\hat{x})(\hat{e}_x)^n = \left( \sum_{j=1}^P \hat{e}_{x|j} \frac{\partial}{\partial x_j} \right)^n F(x) \Big|_{x=\hat{x}},$$

for  $x_j$  and  $\hat{e}_{x|j}$  referring to the  $j$ -th element of  $x$  and  $\hat{e}_x$ . Similarly, expanding  $\bar{x}$  in (11) about  $\hat{x}^-$  and applying Assumption 1 yields

$$\bar{x} = F(\hat{x}^-) + \frac{1}{2}\nabla^2 F(\hat{x}^-)(\hat{e}_x^-)^2 + \cdots. \quad (38)$$

Thus, by subtracting (37) by (38), the estimate error caused by the prior estimate in (11) can be approximated by

$$\bar{e}_x \approx \hat{F}^- \hat{e}_x^- + v^-, \quad (39)$$

where

$$\hat{F}^- = \left( \frac{\partial F(x)}{\partial x} \Big|_{x=\hat{x}^-} \right). \quad (40)$$

However, to accomodate the residuals emerging from this Taylor approximation, we utilise a diagonal matrix defined as  $\alpha = \text{diag}(\alpha_1, \alpha_2, \cdots, \alpha_M)$ , such that we have the approximation formula of error caused by the prior estimate in (11) at time step  $k$  given by

$$\bar{e}_x = \alpha^- \hat{F}^- \hat{e}_x^- + v^-. \quad (41)$$

For measurement approximation, since the measurement estimate is dependent on the prior estimate of  $x$ , the Taylor expansion of  $y$  in (5) is about  $\bar{x}$ . The Taylor expansion of the measurement dynamics at time step  $k$  is given by

$$y = H(\bar{x}) + \nabla H(\bar{x})\bar{e}_x + \frac{1}{2}\nabla^2 H(\bar{x})(\bar{e}_x)^2 + \cdots + w, \quad (42)$$

where

$$\nabla^n H(\bar{x})(\bar{e}_x)^n = \left( \sum_{j=1}^P \bar{e}_{x|j} \frac{\partial}{\partial x_j} \right)^n H(x) \Big|_{x=\bar{x}},$$

for  $x_j$  and  $\bar{e}_{x|j}$  referring to the  $j$ -th element of  $x$  and  $\bar{e}_x$ . Subsequently, expanding the predicted measurement in (15) about  $\bar{x}$  yields

$$\bar{y} = H(\bar{x}) + \frac{1}{2} \nabla^2 H(\bar{x}) (\bar{e}_x)^2 + \dots \quad (43)$$

Hence, subtracting (42) by (43) leads to

$$\bar{e}_y \approx \hat{H} \bar{e}_x + w, \quad (44)$$

where

$$\hat{H} = \left( \frac{\partial H(x)}{\partial x} \bigg|_{x=\bar{x}} \right). \quad (45)$$

An instrumental diagonal matrix related to the measurement prediction added to accomodate the residuals of this approximation is  $\beta = \text{diag}(\beta_1, \beta_2, \dots, \beta_M)$ . Therefore, we have the error of the measurement prediction in (15) formulated as

$$\bar{e}_y = \beta \hat{H} \bar{e}_x + w. \quad (46)$$

In communication approximation, one may follow similar procedures to obtain the approximation of the measurement error. Since the communication estimate is also dependent on the prior state estimate, the expansion of the communication data in (6) is also about  $\bar{x}$ . The Taylor expansion of the communication data at time step  $k$  is

$$z = G(\bar{x}) + \nabla G(\bar{x}) \bar{e}_x + \frac{1}{2} \nabla^2 G(\bar{x}) (\bar{e}_x)^2 + \dots + s, \quad (47)$$

where

$$\nabla^n G(\bar{x}) (\bar{e}_x)^n = \left( \sum_{j=1}^P \bar{e}_{x|j} \frac{\partial}{\partial x_j} \right)^n G(x) \bigg|_{x=\bar{x}},$$

for  $x_j$  and  $\bar{e}_{x|j}$  referring to the  $j$ -th element of  $x$  and  $\bar{e}_x$ . Subsequently, the Taylor expansion of (18) about  $\bar{x}$  is

$$\bar{z} = G(\bar{x}) + \frac{1}{2} \nabla^2 G(\bar{x}) (\bar{e}_x)^2 + \dots \quad (48)$$

Hence, subtracting (47) by (48) yields the approximation error of the predicted communication data in (18) given by

$$\bar{e}_z \approx \hat{G} \bar{e}_x + s, \quad (49)$$

where

$$\hat{G} = \left( \frac{\partial G(x)}{\partial x} \bigg|_{x=\bar{x}} \right). \quad (50)$$

An instrumental diagonal matrix to deal with the residuals of the approximation of this communication data is  $\gamma = \text{diag}(\gamma_1, \gamma_2, \dots, \gamma_M)$ . Hence, the approximate of the communication data satisfies the following equality:

$$\bar{e}_z = \gamma \hat{G} \bar{e}_x + s. \quad (51)$$

The unscented transformation of the covariance also requires us to design a approximation of the estimate covariance matrices. In the case of the prior estimate covariance, we insert (41) to  $\bar{P}_{xx}^* = \text{E}[\bar{e}_x \bar{e}_x^T]$  and use Assumption 2 resulting in

$$\bar{P}_{xx}^* = \alpha^- \hat{F}^- P^- \hat{F}^{-\top} \alpha^{-\top} + Q + \delta_{xx}^- \quad (52)$$

as the approximation of the prior estimate covariance. Parameter  $\delta_{xx}$  is to accommodate the residual approximation error of  $\bar{P}_{xx}$  by  $\bar{P}_{xx}^*$ . For the error covariance matrix of the measurement, with Assumption 2 and  $\bar{e}_y$  from (46), the approximate value is given by  $\bar{P}_{yy}^* = \text{E}[\bar{e}_y \bar{e}_y^T]$ , i.e.,

$$\bar{P}_{yy}^* = \beta \hat{H} \bar{P}_{xx}^* \hat{H}^\top \beta^\top + R + \delta_{yy}, \quad (53)$$

where  $\delta_{yy}$  denotes the approximate error caused by  $\bar{P}_{yy}^*$ . Subsequently, the error covariance matrix of the communication data, the approximation is given by  $\bar{P}_{zz}^* = \text{E}[\bar{e}_z \bar{e}_z^T]$ . Substituting (51) into it and using Assumption 2 yield

$$\bar{P}_{zz}^* = \gamma \hat{G} \bar{P}_{xx}^* \hat{G}^\top \gamma^\top + S + \delta_{zz}, \quad (54)$$

where in this case  $\delta_{zz}$  is the approximate error of  $\bar{P}_{zz}^*$ .

There are three remaining important approximations that have to be defined prior to the performance analysis, i.e., the error covariance between the estimate and measurement error, between the estimate and communication error, and between the measurement and communication error. These values are  $\bar{P}_{xy}^* = \text{E}[\bar{e}_x \bar{e}_y^T]$ ,  $\bar{P}_{xz}^* = \text{E}[\bar{e}_x \bar{e}_z^T]$  and  $\bar{P}_{yz}^* = \text{E}[\bar{e}_y \bar{e}_z^T]$ . With (41), (46) and (51), we have

$$\bar{P}_{xy}^* = \bar{P}_{xx}^* \hat{H}^\top \beta^\top + \delta_{xy}, \quad (55)$$



$$\bar{P}_{xz}^* = \bar{P}_{xx}^* \hat{G}^\top \gamma^\top + \delta_{xz}, \quad (56)$$

and

$$\bar{P}_{yz}^* = \beta \hat{H} \bar{P}_{xx}^* \hat{G}^\top \gamma^\top + \delta_{yz}, \quad (57)$$

by denoting  $\delta_{xy}$ ,  $\delta_{xz}$  and  $\delta_{yz}$  as the approximation errors caused by  $\bar{P}_{xy}^*$ ,  $\bar{P}_{xz}^*$  and  $\bar{P}_{yz}^*$ , respectively. Notice that, in above three equations, we have also applied the conditions in Assumption 2.

#### 4.4. Performance Analysis

The following assumptions are made prior to analysing the performance of the proposed filter.

**Assumption 5.** *There exist non-zero real numbers  $\underline{\alpha}$ ,  $\underline{\beta}$ ,  $\underline{\gamma}$ ,  $\underline{f}$ ,  $\underline{h}$ ,  $\underline{g}$ , and  $\bar{\alpha}$ ,  $\bar{\beta}$ ,  $\bar{\gamma}$ ,  $\bar{f}$ ,  $\bar{h}$ ,  $\bar{g}$ , for every  $k > 0$ ,  $k \in \mathbb{Z}^*$ , such that*

$$\underline{\alpha}^2 I \leq \alpha \alpha^T \leq \bar{\alpha}^2 I, \quad (58)$$

$$\underline{\beta}^2 I \leq \beta \beta^T \leq \bar{\beta}^2 I, \quad (59)$$

$$\underline{\gamma}^2 I \leq \gamma \gamma^T \leq \bar{\gamma}^2 I, \quad (60)$$

$$\underline{f}^2 I \leq \hat{F} \hat{F}^T \leq \bar{f}^2 I, \quad (61)$$

$$\underline{h}^2 I \leq \hat{H} \hat{H}^T \leq \bar{h}^2 I. \quad (62)$$

$$\underline{g}^2 I \leq \hat{G} \hat{G}^T \leq \bar{g}^2 I. \quad (63)$$

**Assumption 6.** *There exist positive real numbers  $\underline{p}$ ,  $\underline{q}$ ,  $\underline{r}$ ,  $\underline{s}$ ,  $\underline{\delta}_{xx}$ ,  $\underline{\delta}_{yy}$ ,  $\underline{\delta}_{zz}$ ,  $\underline{\delta}_{xy}$ ,  $\underline{\delta}_{xz}$ ,  $\underline{\delta}_{yz}$ , and  $\bar{p}$ ,  $\bar{q}$ ,  $\bar{r}$ ,  $\bar{s}$ ,  $\bar{\delta}_{xx}$ ,  $\bar{\delta}_{yy}$ ,  $\bar{\delta}_{zz}$ ,  $\bar{\delta}_{xy}$ ,  $\bar{\delta}_{xz}$ ,  $\bar{\delta}_{yz}$ , for every  $k > 0$ ,  $k \in \mathbb{Z}^*$ , such that*

$$\underline{p} I \leq P \leq \bar{p} I, \quad (64)$$

$$\underline{q} I \leq Q \leq \bar{q} I, \quad (65)$$

$$\underline{r} I \leq R \leq \bar{r} I, \quad (66)$$

$$\underline{s} I \leq S \leq \bar{s} I, \quad (67)$$

$$\underline{\delta}_{xx} I \leq \delta_{xx} \leq \bar{\delta}_{xx} I, \quad (68)$$

$$\underline{\delta}_{yy} I \leq \delta_{yy} \leq \bar{\delta}_{yy} I, \quad (69)$$

$$\underline{\delta}_{zz}I \leq \delta_{zz} \leq \bar{\delta}_{zz}I, \quad (70)$$

$$\underline{\delta}_{xy}I \leq \delta_{xy} \leq \bar{\delta}_{xy}I, \quad (71)$$

$$\underline{\delta}_{xz}I \leq \delta_{xz} \leq \bar{\delta}_{xz}I, \quad (72)$$

$$\underline{\delta}_{yz}I \leq \delta_{yz} \leq \bar{\delta}_{yz}I, \quad (73)$$

Some lemmas stating the upper and lower boundaries of some parameters are declared as follows.

**Lemma 2.** *Let the conditions in Assumptions 5 and 6 holds. Let the error of the prior estimates of the nonlinear Kalman filter be approximated using (41), (46) and (51). With conditions in Assumptions 5 and 6 hold, the prior covariance matrices are bounded by*

$$\underline{\alpha}^2 \underline{f}^2 \underline{p} + \underline{q} + \underline{\delta}_{xx} \leq \bar{P}_{xx}^* \leq \bar{\alpha}^2 \bar{f}^2 \bar{p} + \bar{q} + \bar{\delta}_{xx}, \quad (74)$$

$$\underline{\beta}^2 \underline{h}^2 (\underline{\alpha}^2 \underline{f}^2 \underline{p} + \underline{q} + \underline{\delta}_{xx}) + \underline{r} + \underline{\delta}_{yy} \leq \bar{P}_{yy}^* \leq \bar{\beta}^2 \bar{h}^2 (\bar{\alpha}^2 \bar{f}^2 \bar{p} + \bar{q} + \bar{\delta}_{xx}) + \bar{r} + \bar{\delta}_{yy}, \quad (75)$$

$$\underline{\gamma}^2 \underline{g}^2 (\underline{\alpha}^2 \underline{f}^2 \underline{p} + \underline{q} + \underline{\delta}_{xx}) + \underline{s} + \underline{\delta}_{zz} \leq \bar{P}_{zz}^* \leq \bar{\gamma}^2 \bar{g}^2 (\bar{\alpha}^2 \bar{f}^2 \bar{p} + \bar{q} + \bar{\delta}_{xx}) + \bar{s} + \bar{\delta}_{zz}, \quad (76)$$

$$(\underline{\alpha}^2 \underline{f}^2 \underline{p} + \underline{q} + \underline{\delta}_{xx}) \hat{H}^\top \beta^\top + \underline{\delta}_{xy} \leq \bar{P}_{xy}^* \leq (\bar{\alpha}^2 \bar{f}^2 \bar{p} + \bar{q} + \bar{\delta}_{xx}) \hat{H}^\top \beta^\top + \bar{\delta}_{xy}, \quad (77)$$

$$(\underline{\alpha}^2 \underline{f}^2 \underline{p} + \underline{q} + \underline{\delta}_{xx}) \hat{G}^\top \gamma^\top + \underline{\delta}_{xz} \leq \bar{P}_{xz}^* \leq (\bar{\alpha}^2 \bar{f}^2 \bar{p} + \bar{q} + \bar{\delta}_{xx}) \hat{G}^\top \gamma^\top + \bar{\delta}_{xz}, \quad (78)$$

$$(\underline{\alpha}^2 \underline{f}^2 \underline{p} + \underline{q} + \underline{\delta}_{xx}) \beta \hat{H} \hat{G}^\top \gamma^\top + \underline{\delta}_{yz} \leq \bar{P}_{yz}^* \leq (\bar{\alpha}^2 \bar{f}^2 \bar{p} + \bar{q} + \bar{\delta}_{xx}) \beta \hat{H} \hat{G}^\top \gamma^\top + \bar{\delta}_{yz}. \quad (79)$$

**Lemma 3.** *Let the error of the prior estimates of the nonlinear Kalman filter be approximated using (41), (46) and (51); and the prior error covariances bounded by (74)-(79). Then, with the conditions in Assumptions 5 and 6 hold, the square of communication gain is bounded by*

$$\underline{c}^2 I \leq CC^\top \leq \bar{c}^2 I \quad (80)$$

where

$$\underline{c}^2 = \frac{\underline{\gamma}^2 \underline{g}^2 (\underline{\alpha}^2 \underline{f}^2 \underline{p} + \underline{q} + \underline{\delta}_{xx})^2 + \underline{\delta}_{xz}^2}{(\bar{\gamma}^2 \bar{g}^2 (\bar{\alpha}^2 \bar{f}^2 \bar{p} + \bar{q} + \bar{\delta}_{xx}) + \bar{s} + \bar{\delta}_{zz})^2} \quad (81)$$

and

$$\bar{c}^2 = \frac{2 \left( \bar{\gamma}^2 \bar{g}^2 (\bar{\alpha}^2 \bar{f}^2 \bar{p} + \bar{q} + \bar{\delta}_{xx})^2 + \bar{\delta}_{xz}^2 \right)}{(\underline{\gamma}^2 \underline{g}^2 (\underline{\alpha}^2 \underline{f}^2 \underline{p} + \underline{q} + \underline{\delta}_{xx}) + \underline{r} + \underline{\delta}_{zz})^2}. \quad (82)$$

**Lemma 4.** *Let the error of the prior estimates of the nonlinear Kalman filter be approximated using (41), (46) and (51); and the prior error covariances bounded by (74)-(79). Then, with the conditions in Assumptions 5 and 6 hold, the square of Kalman gain is bounded by*

$$\underline{k}^2 I \leqslant K K^\top \leqslant \bar{k}^2 I \quad (83)$$

where

$$\underline{k}^2 = 0 \quad (84)$$

and

$$\bar{k}^2 = \frac{4 \left( (\bar{\beta}^2 \bar{h}^2 + \bar{c}^2 \bar{\beta}^2 \bar{\gamma}^2 \bar{h}^2 \bar{g}^2) (\bar{\alpha}^2 \bar{f}^2 \bar{p} + \bar{q} + \bar{\delta}_{xx})^2 + \bar{\delta}_{xy}^2 + \bar{\delta}_{yz}^2 \right)}{(\underline{\beta}^2 \underline{h}^2 (\underline{\alpha}^2 \underline{f}^2 \underline{p} + \underline{q} + \underline{\delta}_{xx}) + \underline{r} + \underline{\delta}_{yy})^2}, \quad (85)$$

with  $\bar{c}^2$  as the upper bound of the square of communication gain.

**Lemma 5.** *Let  $K$  in (29) and  $C$  in (30) be bounded by (83) and (80), respectively. Define  $M = I - K\beta\hat{H} - C\gamma\hat{G}$ . Then, with the conditions in Assumptions 5 and 6 satisfied, the boundaries of  $MM^\top$  are given by*

$$\underline{m}^2 I \leqslant M M^\top \leqslant \bar{m}^2 I \quad (86)$$

where

$$\underline{m}^2 = 0, \quad (87)$$

and

$$\bar{m}^2 = 3 \left( 1 + \bar{k}^2 \bar{\beta}^2 \bar{h}^2 + \bar{c}^2 \bar{\gamma}^2 \bar{g}^2 \right). \quad (88)$$

**Lemma 6.** *Consider the posterior estimate and covariance update of the distributed nonlinear Kalman filter formulated by (23) and (31), respectively. Define a matrix  $M = (I - K\beta\hat{H} - C\gamma\hat{G})$ . If the conditions in Assumptions 5 and 6 hold, then, at every time step  $k \in \mathbb{Z}^*$ , the inverse of the covariance, denoted by  $\Pi = P^{-1}$ , satisfies the following condition*

$$\Pi \leqslant \varphi \Pi^-, \quad (89)$$

where

$$\varphi = \frac{1}{\underline{\alpha}^2 \underline{f}^2 \underline{m}^2} \left( 1 + \frac{\underline{m}^2 \underline{q} + \underline{k}^2 \underline{r} + \underline{c}^2 \underline{s} + \underline{m}^2 \underline{\delta}_{xx}}{\bar{\alpha}^2 \bar{f}^2 \bar{m}^2 \bar{p}} \right)^{-1}. \quad (90)$$

*Proof.* The error of the posterior estimate is  $\hat{e}_x = x - \hat{x}$  which, by plugging (23) into it, gives

$$\hat{e}_x = \bar{e}_x - K\bar{e}_y - C\bar{e}_z \quad (91)$$

where  $\bar{e}_x = x - \bar{x}$ ,  $\bar{e}_y = y - \bar{y}$  and  $\bar{e}_z = z - \bar{z}$ . The measurement and communication errors can be substituted by the approximate values in (46) and (51), respectively. From this substitution, we have

$$\hat{e}_x = M\bar{e}_x - Kw - Cs \quad (92)$$

where  $M = (I - K\beta\hat{H} - C\gamma\hat{G})$ . Recall the definition of posterior covariance matrix  $P = E[\hat{e}_x\hat{e}_x^\top]$ . Plugging (92) into it gives

$$P = M\bar{P}_{xx}^*M^\top + K RK^\top + CSC^\top \quad (93)$$

in which Assumptions 2 and 3 have been applied. Notice that (93) uses the approximate value of  $\bar{P}_{xx}$ . Using (52), one may have (93) expanded to

$$P = M \left( \alpha^- \hat{F}^- P^- (\alpha^- \hat{F}^-)^\top + Q + \delta_{xx}^- \right) M^\top + K RK^\top + CSC^\top. \quad (94)$$

Rearranging the terms, we have

$$\begin{aligned} P = & M \alpha^- \hat{F}^- \left( I + (M \alpha^- \hat{F}^-)^{-1} K RK^\top (M \alpha^- \hat{F}^-)^{-\top} (P^-)^{-1} \right. \\ & + (M \alpha^- \hat{F}^-)^{-1} C S C^\top (M \alpha^- \hat{F}^-)^{-\top} (P^-)^{-1} \\ & + (\alpha^- \hat{F}^-)^{-1} Q (\alpha^- \hat{F}^-)^{-\top} (P^-)^{-1} \\ & \left. + (\alpha^- \hat{F}^-)^{-1} \delta_{xx}^- (\alpha^- \hat{F}^-)^{-\top} (P^-)^{-1} \right) P^- (M \alpha^- \hat{F}^-)^\top. \end{aligned} \quad (95)$$

Using Assumptions 5 and 6 along with Lemmas 3, 4 and 5, (95) can be expressed as

$$P \geq \underline{\alpha}^2 \underline{f}^2 \underline{m}^2 \left( 1 + \frac{\underline{m}^2 \underline{q} + \underline{k}^2 \underline{r} + \underline{c}^2 \underline{s} + \underline{m}^2 \underline{\delta}_{xx}}{\underline{\alpha}^2 \underline{f}^2 \underline{m}^2 \underline{p}} \right) P^-. \quad (96)$$

Therefore, with  $\Pi = P^{-1}$ , taking the inverse of (96) results in the inverse covariance matrix bounded by

$$\Pi \leq \varphi \Pi^-, \quad (97)$$

where

$$\varphi = \frac{1}{\underline{\alpha}^2 \underline{f}^2 \underline{m}^2} \left( 1 + \frac{\underline{m}^2 \underline{q} + \underline{k}^2 \underline{r} + \underline{c}^2 \underline{s} + \underline{m}^2 \underline{\delta}_{xx}}{\underline{\alpha}^2 \underline{f}^2 \underline{m}^2 \underline{p}} \right)^{-1}. \quad (98)$$

This completes the proof.  $\square$

The following theorem guarantees the boundedness of the proposed distributed nonlinear Kalman filter.

**Theorem 2** (Stochastic Stability of Distributed Nonlinear Kalman Filter). *Let the posterior estimator of the distributed nonlinear Kalman filter be given by (23), and the covariance by (31). Suppose that Assumptions 1 to 6 hold. Then, there exists a stochastic function  $V \geq 0$  such that the estimate error is bounded in mean square exponentially.*

*Proof.* To proof the theorem, first, let us choose a stochastic function candidate

$$V = \hat{e}_x^\top \Pi \hat{e}_x \quad (99)$$

where  $\hat{e}_x = x - \hat{x}$  and  $\Pi = P^{-1}$ . To show the boundedness of  $V$ , multiplying the inequality in the condition (64) by  $\hat{e}_x^\top \hat{e}_x$  results in

$$\frac{1}{\underline{p}} \|\hat{e}_x\|^2 \leq V(\hat{e}_x) \leq \frac{1}{\underline{p}} \|\hat{e}_x\|^2 \quad (100)$$

$$\underline{v} \|\hat{e}_x\|^2 \leq V(\hat{e}_x) \leq \bar{v} \|\hat{e}_x\|^2. \quad (101)$$

Therefore, since  $\underline{p}$  and  $\bar{p}$  are positive,  $V$  is positively bounded with  $\underline{v}, \bar{v} > 0$ . Since  $V$  is positively bounded, it can now be used to show that there exist  $\mu > 0$  and  $0 < \sigma \leq 1$  such that the following inequality holds

$$\mathbb{E}[V^+(\hat{e}_x^+)|\hat{e}_x] \leq \mu + (1 - \sigma)V(\hat{e}_x). \quad (102)$$

Let us consider the stochastic function (101) at time step  $k + 1$ , i.e.,

$$V^+(\hat{e}_x^+) = (\hat{e}_x^+)^{\top} \Pi^+ \hat{e}_x^+. \quad (103)$$

Using the definition of the errors for the estimates, i.e.,  $\bar{e}_x^+ = x^+ - \bar{x}^+$ ,  $\bar{e}_y^+ = y^+ - \bar{y}^+$ , and  $\bar{e}_z^+ = z^+ - \bar{z}^+$ , and plugging them into the posterior estimate (23) at time  $k + 1$  yields the posterior error  $\hat{e}_x^+ = x^+ - \hat{x}^+$  given by

$$\hat{e}_x^+ = \bar{e}_x^+ - K^+ \bar{e}_y^+ - C \bar{e}_z^+. \quad (104)$$

By inserting (41), (46) and (51) at  $k + 1$  to (104), we have

$$\hat{e}_x^+ = M^+ \alpha \hat{F} \hat{e}_x + M^+ v - K^+ w^+ - C^+ s^+ \quad (105)$$

where  $M^+ = (I - K^+ \beta^+ \hat{H}^+ - C^+ \gamma^+ \hat{G}^+)$ .

Plugging (105) into (103) and taking its expected value with respect to  $\hat{e}_x$  lead to

$$\begin{aligned} \mathbb{E}[V^+(\hat{e}_x^+)|\hat{e}_x] &= \mathbb{E}[(M^+ \alpha \hat{F} \hat{e}_x)^\top \Pi^+ M^+ \alpha \hat{F} \hat{e}_x + (M^+ v)^\top \Pi^+ M^+ v \\ &\quad + (K^+ w^+)^\top \Pi^+ K^+ w^+ + (C^+ s^+)^\top \Pi^+ C^+ s^+ | \hat{e}_x] \end{aligned} \quad (106)$$

where Assumptions 3 and 4 have been utilised. With Assumptions 5, applying Lemma 5 and Lemma 6 to the first term of (106) leads to

$$\begin{aligned} \mathbb{E}[(M^+ \alpha \hat{F} \hat{e}_x)^\top \Pi_{xx}^+ M^+ \alpha \hat{F} \hat{e}_x | \hat{e}_x] &\leq \mathbb{E}[\text{tr}(\bar{\alpha}^2 \bar{f}^2 \bar{m}^2 \hat{e}_x^\top \Pi_{xx}^+ \hat{e}_x) | \hat{e}_x] \\ &\leq \bar{\alpha}^2 \bar{f}^2 \bar{m}^2 \varphi V(\hat{e}_x) \end{aligned} \quad (107)$$

in which we have applied the properties of matrix trace  $\text{tr}(\mathbb{E}[A]) = \mathbb{E}[\text{tr}(A)]$  and  $\text{tr}(AB) = \text{tr}(BA)$ . Similarly, employing Assumptions 1, 5, and 6, along with Lemmas 3, 4, and 5 to the second, third and fourth terms of (107) give

$$\begin{aligned} \mathbb{E}[(M^+ v)^\top \Pi_{xx}^+ M^+ v + (K^+ w^+)^\top \Pi_{xx}^+ K^+ w^+ + (C^+ s^+)^\top \Pi_{xx}^+ C^+ s^+ | \hat{e}_x] \\ \leq \frac{1}{\underline{p}} \left( \bar{m}^2 \bar{q} + \bar{k}^2 \bar{r} + \bar{c}^2 \bar{s} \right) \end{aligned} \quad (108)$$

where the properties of a trace of matrix have also been used. By denoting

$$\mu = \frac{1}{\underline{p}} \left( \bar{m}^2 \bar{q} + \bar{k}^2 \bar{r} + \bar{c}^2 \bar{s} \right), \quad (109)$$

$$\sigma = 1 - \bar{\alpha}^2 \bar{f}^2 \bar{m}^2 \varphi, \quad (110)$$

and choosing some constants such that  $0 \leq \sigma \leq 1$  and  $\mu \geq 0$ , we are now able to write (108) as

$$\mathbb{E}[V^+(\hat{e}_x^+)|\hat{e}_x] \leq \mu + (1 - \sigma) V(\hat{e}_x). \quad (111)$$

Therefore, by employing Lemma 1, we can conclude that  $\hat{e}_x$  is exponentially bounded in mean square with

$$\mathbb{E}[\|\hat{e}_x\|^2] \leq \frac{\bar{v}}{\underline{v}} \mathbb{E}[\|\hat{e}_x(0)\|^2] (1 - \sigma)^k + \frac{\mu}{\underline{v}} \sum_{l=1}^{k-1} (1 - \sigma)^l. \quad (112)$$

This completes the proof.  $\square$

---

**Algorithm 1** *Distributed Nonlinear Kalman Filter*

---

Input:  $x(0), Q(0), R(0), S(0), F(\cdot), H(\cdot), G(\cdot)$ .

Return:  $\hat{x}, P$ .

If  $k = 0$ , execute

*Initialisation:*

- 1: Initialise the estimator  $\hat{x}$  and covariance  $P$ :

$$\tilde{x} = [\hat{x}(0)^\top \ 0_p \ 0_q \ 0_r]^\top, \text{ for } \hat{x}(0) = E[x(0)],$$

$$\tilde{P} = \text{diag}(P(0), Q(0), R(0), S(0)), \text{ for } P(0) = E[(x(0) - \hat{x}(0))(x(0) - \hat{x}(0))^\top].$$

If  $k = 1, 2, \dots, T_k, T_k < \infty$ , execute

*Prediction:*

- 1: Calculate the sigma vector  $\mathcal{X}^-$  using Eq. (9).
- 2: Propagate the sigma vector through  $F(\cdot), H(\cdot), G(\cdot)$  using (10), (14) and (17), respectively.
- 3: Calculate the prior estimate  $\bar{x}$  using (11),  $\bar{y}_i$  using (15), and  $\bar{z}$  using (18).
- 4: Calculate the prior covariances  $\bar{P}_{xx}, \bar{P}_{yy}, \bar{P}_{zz}, \bar{P}_{xy}, \bar{P}_{xz}$ , and  $\bar{P}_{yz}$  using Eqs. (12), (16), (19), (20), (21), and (22), respectively.

*Update:*

Each agent exchanges its predicted value to its neighboring agents.

Update the Kalman and communication gains,  $K$  and  $C$  using Eqs. (29) and (30), respectively.

Update the posterior estimate  $\hat{x}$  using Eq. (23).

Update the posterior covariance  $P$  using Eq. (31)).

- 1: Update the Kalman and communication gains,  $K$  and  $C$  using Eqs. (29) and (30), respectively.
  - 2: Update the posterior estimate  $\hat{x}$  using Eq. (23).
  - 3: Update the posterior covariance  $P$  using Eq. (31)).
- 

#### 4.5. Practical Algorithm

This proposed algorithm can generally be applied to solve many estimation problem in, but not limited to, engineering field via distributed computation of the filter whenever the dynamical process of a system, measurement and communication update can be formulated as in the discussion. For example, this algorithm can further be employed to estimate the Bellman value and Q function in reinforcement learning, such as solving the learning problem in [30] in a distributed

manner using our proposed filtering algorithm. In the case of sensor network, this algorithm can also be employed to estimate a spatial information of an area, which will be discussed in the next section.

The procedure of implementing the proposed distributed nonlinear Kalman filter is summarised in Algorithm 1. In the beginning, to implement the algorithm in a distributed manner, one should supply some initial values of the augmented state variable and the augmented covariance matrix in each agent. Afterwards, once the system starts, in every iteration each agent will locally perform the prediction step based on the latest available estimate value and its covariances using unscented transformation technique. This prediction step produces the prior estimate values and their covariance and cross-covariance matrices. The final stage of each iteration is called the update stage. In this stage, each agent will have to exchange its local prior estimate values and their covariance matrices to its neighboring agents based on the network topology. This step is followed by calculating the new posterior estimate value and its covariance matrix. These prediction and update procedures repeat iteratively.

## 5. Application to Distributed Coverage Control

In this section, we implement the proposed distributed nonlinear Kalman filter algorithm to specifically solve an existing problem in sensor network. The problem considered in this section is the field estimation of coverage control problem. This part is extended from our previous work in distributed coverage control with known information distribution elaborated in [22].

### 5.1. Locational Optimisation

Consider  $n$  robots deployed in a convex set  $\mathcal{Q} \in \mathbb{R}^d$ , and its connection topology represented by graph  $\mathcal{G}_n(\mathcal{V}_n, \mathcal{E}_n)$ . The Laplacian of this graph is denoted by  $\mathcal{L}_n$ . The set containing the position of robots is denoted by  $\mathcal{P} = \{p_1, p_2, \dots, p_n\}$ . In the real environment, the sensing quality of  $i$ -th sensor might be affected by noises around it. Therefore, we need to design a controller that is able to drive the robot towards the optimal position to acquire the physical quantity at the point  $q \in \mathcal{Q}$ .

Sensing unreliability function  $\rho(\|q - p_i\|)$ ,  $\rho : \mathbb{R}^+ \rightarrow \mathbb{R}^+$ , denotes a function that provides the quantitative information of the sensing performance at point  $q$  monitored by agent  $i$  at  $p_i$ . We assume the sensing unreliability function to be isotropic, increasing and convex. In the sense of this unreliability function, the isotropic, increasing and convex sensing function means the sensing quality at



point  $q$  sensed by robot  $i$  decreases proportionally with the distance  $\|q - p_i\|$  – the greater is the value of sensing unreliability function, the worse is the sensing quality.

Density function, or information distribution function, is a measure of information probability in an area. It indicates the importance of a quantity to measure at the particular point  $q \in \mathcal{Q}$ . In this case, density function is denoted as  $\phi(q)$ ,  $\phi : \mathcal{Q} \mapsto \mathbb{R}_+$ .

After the sensing unreliability function and density function are described, we are now ready to introduce the locational optimisation problem. Notice that, having  $\mathcal{P}$  as the network configuration, we are able to use the Voronoi partition of  $\mathcal{Q}$  described as

$$\mathcal{R}_i = \{q \in \mathcal{Q} \mid \|q - p_i\| \leq \|q - p_j\|, \forall i \in V_n, j \neq i\}.$$

Thus, locational optimisation is defined as a problem aiming to find the best location of a group of agents within an area according to a particular cost function. With the given descriptions, the corresponding cost function of locational optimisation is formulated as

$$\mathcal{H}(\mathcal{P}) = \sum_{i=1}^n \int_{\mathcal{R}_i} \rho(\|q - p_i\|) \phi(q) dq. \quad (113)$$

It can be seen that the best locations of the sensor are achieved if  $\mathcal{H}$  is minimized. Under the convexity assumption of the sensing unreliability function, by using the properties of convex function in [31], it can be concluded that the objective function is convex in  $\mathcal{P}$ .

Furthermore, by employing the concepts of rigid body, we have the mass, moment of inertia and centroid of  $i$ -th Voronoi region expressed as

$$M_{\mathcal{R}_i} = \int_{\mathcal{R}_i} \phi(q) dq, \quad \mathcal{I}_{\mathcal{R}_i} = \int_{\mathcal{R}_i} q \phi(q) dq, \quad \text{and} \quad C_{\mathcal{R}_i} = \frac{\mathcal{I}_{\mathcal{R}_i}}{M_{\mathcal{R}_i}}, \quad (114)$$

respectively. Therefore, applying the parallel-axis theorem and choosing  $\rho(\|q - p_i\|) = \|q - p_i\|^2$ , the cost function in (113) can equivalently be formulated as

$$\mathcal{H}(\mathcal{P}) = \sum_{i=1}^n \mathcal{H}_{\mathcal{R}_i} + \sum_{i=1}^n M_{\mathcal{R}_i} \|p_i - C_{\mathcal{R}_i}\|^2, \quad (115)$$

which is also convex. The augmented expression of (115) over the network, we may write

$$\mathcal{H}(\mathcal{P}) = \mathcal{H}_{\mathcal{R}} + (p - C_{\mathcal{R}})^\top M_{\mathcal{R}} (p - C_{\mathcal{R}}), \quad (116)$$

where  $\mathcal{H}_{\mathcal{R}} = \sum_{i=1}^n \mathcal{H}_{\mathcal{R}_i}$ ,  $p = \text{vec}_n(p_i)$ ,  $C_{\mathcal{R}} = \text{vec}_n(C_{\mathcal{R}_i})$ .

Since the Delaunay triangulation is applied to generate Voronoi partition without limited communication range, in this coverage scenario we will always have an undirected and connected Delaunay graph of the network.

## 5.2. Distributed Optimisation

In this part, the coverage controller from distributed optimisation point is briefly elaborated. Let the augmented vector of the position of robots be defined as  $p = \text{vec}_n(p_i)^\top \in \mathbb{R}^{nd}$ , for  $i \in \mathcal{V}_n$ , and the augmented Laplacian matrix of the Delaunay graph as  $\hat{\mathcal{L}} = \mathcal{L} \otimes I_d \in \mathbb{R}^{nd \times nd}$ . From our previous result in [22], the distributed expression of cost function of coverage problem (113) can be formulated as

$$\min_{p_i} \quad \mathcal{H}(\mathcal{P}) = \sum_{i=1}^n \int_{\mathcal{R}_i} \rho(\|q - p_i\|) \phi(q) dq, \quad (117a)$$

$$\text{s.t.} \quad \hat{\mathcal{L}}(p - C_{\mathcal{R}}) = 0, \quad (117b)$$

taking  $C_{\mathcal{R}} = \text{vec}_n(C_{\mathcal{R}_i}) \in \mathbb{R}^{nd}$ , for  $i \in \mathcal{V}_n$ , as the offset value.

Subsequently, we augment the constraint to the objective function using the Lagrange multiplier method. The Lagrangian of this coverage problem,  $\tilde{\mathcal{H}} : \mathbb{R}^{nd} \times \mathbb{R}^{nd} \rightarrow \mathbb{R}$ , with disagreement function of consensus protocol is

$$\tilde{\mathcal{H}}(p, \nu) = \mathcal{H} + \nu^\top \hat{\mathcal{L}}(p - C_{\mathcal{R}}) + \frac{1}{2}(p - C_{\mathcal{R}})^\top \hat{\mathcal{L}}(p - C_{\mathcal{R}}), \quad (118)$$

where  $\nu \in \mathbb{R}^{nd}$  defines the Lagrangian multiplier of the constraint. Thus, the control input and Lagrangian multiplier update, which is derived by taking the gradient of  $\tilde{\mathcal{H}}(p, \nu)$  with respect to  $p$  and  $\nu$ , respectively, can be expressed as

$$u = -2\alpha M_{\mathcal{R}}(p - C_{\mathcal{R}}) - \beta \hat{\mathcal{L}}(p - C_{\mathcal{R}}) - \gamma \hat{\mathcal{L}}\nu, \quad (119a)$$

$$\dot{\nu} = \eta \hat{\mathcal{L}}(p - C_{\mathcal{R}}), \quad (119b)$$

where  $\alpha, \beta, \gamma, \eta \in \mathbb{R}^+$ .

To apply the estimation update from the online field estimator,  $C_{\mathcal{R}_i}$  in the control input is modified to be  $\hat{C}_{\mathcal{R}_i}$  such that the control input of agent  $i$  is given by

$$u = -k_p(p - \hat{C}_{\mathcal{R}}) - k_c \hat{\mathcal{L}}(p - \hat{C}_{\mathcal{R}}) - k_i \hat{\mathcal{L}} \int_{t_a}^{t_b} (p - \hat{C}_{\mathcal{R}}) dt. \quad (120)$$

Notice that, in the above expression, we have simplified the gains expression by choosing  $k_p, k_c, k_i \in \mathbb{R}^+$  to be constant gains.

### 5.3. Field Estimation

In this estimation scenario, we consider the information distribution of an area can be expressed as

$$\phi(q) = \mathcal{K}(q)^\top \theta, \quad (121)$$

with  $\mathcal{K} : \mathcal{Q} \mapsto \mathbb{R}^m$  denoting a vector of basis function that is known by each robot [25] [32]. Weight  $\theta \in \mathbb{R}^m$  here denotes an unknown parameter to estimate. From (114), it can be seen that the mass  $M_{\mathcal{R}_i}$  of the centroid  $C_{\mathcal{R}_i}$  must be greater than zero. It also implies that  $\phi(q) > 0$ . Therefore, a boundary must be set, i.e.,

$$\theta_l \geq \underline{\theta}, \quad \text{for } l \in \{1, 2, \dots, m\}, \quad (122)$$

with  $\underline{\theta} > 0$  denoting the lower bound of  $l$ -th entry of  $\theta$ .

Afterwards, several estimation variables are defined.  $\hat{\theta}_i$  and  $\hat{\phi}_i$  denote, respectively, the estimate weight and information distribution of robot  $i$ . Let  $\hat{\phi}_i = \mathcal{K}^\top \hat{\theta}_i$  denote the estimate of  $\phi$  of  $i$ -th robot. Then, we may have the estimate of mass, moment of inertia and centroid of Voronoi region formulated as

$$\hat{M}_{\mathcal{R}_i} = \int_{\mathcal{R}_i} \hat{\phi}_i(q) dq, \quad \hat{\mathcal{I}}_{\mathcal{R}_i} = \int_{\mathcal{R}_i} q \hat{\phi}_i(q) dq, \quad \text{and} \quad \hat{C}_{\mathcal{R}_i} = \frac{\hat{\mathcal{I}}_{\mathcal{R}_i}}{\hat{M}_{\mathcal{R}_i}}, \quad (123)$$

respectively.

By letting  $\phi_i$  be the measurement value read by sensor  $i$ , the estimation error of  $\theta$  and  $\phi$  can, respectively, be expressed as  $\tilde{\theta}_i = \theta - \hat{\theta}_i$  and  $\tilde{\phi}_i = \phi_i - \hat{\phi}_i = \mathcal{K}(q)^\top \tilde{\theta}_i$ . With these notations, we have the error between true and estimate value of the rigid body properties in the form of

$$\tilde{M}_{\mathcal{R}_i} = \int_{\mathcal{R}_i} \mathcal{K}^\top(q) \tilde{\theta}_i dq, \quad \tilde{\mathcal{I}}_{\mathcal{R}_i} = \int_{\mathcal{R}_i} q \mathcal{K}^\top(q) \tilde{\theta}_i dq, \quad (124)$$

and

$$\tilde{C}_{\mathcal{R}_i} = \frac{\tilde{\mathcal{I}}_{\mathcal{R}_i}}{\tilde{M}_{\mathcal{R}_i}}. \quad (125)$$

Therefore, the distributed field estimation of the coverage control problem is solved if the following conditions are satisfied:

1.  $\lim_{k \rightarrow \infty} \|p_i - \hat{C}_{\mathcal{R}_i}\| = 0$ , for  $i \in \mathcal{V}_n$ ,
2.  $\lim_{k \rightarrow \infty} \mathcal{K}^\top \tilde{\theta}_i = 0$ , for  $i \in \mathcal{V}_n$ .

It can also be concluded that if  $\hat{\phi}_i = \phi$ , then  $\hat{M}_{\mathcal{R}_i} = M_{\mathcal{R}_i}$ ,  $\hat{\mathcal{I}}_{\mathcal{R}_i} = \mathcal{I}_{\mathcal{R}_i}$  and  $\hat{C}_{\mathcal{R}_i} = C_{\mathcal{R}_i}$ .

#### 5.4. Distributed Estimation with Laplacian-based Graph

To obtain the estimation of the centroid required in the control law (120), the distributed nonlinear Kalman filter we have designed in Section 4 is utilised.

Consider the estimation dynamics, consisting of the weight of the density function, sensor reading and communication dynamics which, in augmented expression, are given by

$$\theta^+ = \theta + v, \quad (126)$$

$$\psi = \phi(q, \theta) + w, \quad (127)$$

$$\tau = \hat{\mathcal{L}}\theta + s, \quad (128)$$

respectively. Note that  $\theta = \text{vec}_n(\theta_i)$ ,  $\psi = \text{vec}_n(\psi_i)$ ,  $\tau = \text{vec}_n(\tau_i)$ ,  $v = \text{vec}_n(v_i)$ ,  $w = \text{vec}_n(w_i)$ ,  $s = \text{vec}_n(s_i)$  and  $\hat{\mathcal{L}} = \mathcal{L}_n \otimes I_p$ .

Since in this application the network topology is modelled using the Laplacian matrix, the posterior estimator in (23) may be rewritten as

$$\hat{x} = \bar{x} + K(y - \bar{y}) - C\hat{\mathcal{L}}(\bar{x} - x). \quad (129)$$

Let the element-wise expression of (129) be given by

$$\hat{x}_i = \bar{x}_i + K_i(y_i - \bar{y}_i) + C_i \sum_{j \in N_i} ((\bar{x}_j - x_j) - (\bar{x}_i - x_i)), \text{ for } i \in \mathcal{V}_n. \quad (130)$$

Since each agent is assigned to track the same dynamical process, we have  $x_i = x_j = x$ . Thus, the posterior estimator of the Laplacian-based communication topology can be reduced to

$$\hat{x}_i = \bar{x}_i + K_i(y_i - \bar{y}_i) + C_i \sum_{j \in N_i} (\bar{x}_j - \bar{x}_i), \text{ for } i \in \mathcal{V}_n. \quad (131)$$

The coverage control procedure with the Kalman-based estimator is summarised in Algorithm 2.

## 6. Numerical Experiments

For comparison, two simulation scenarios are presented to estimate the spatial information distribution of an area and drive a group of robots to estimate the optimal configurations using two different estimation algorithms. The first scenario, referred to as the modified-consensus observer algorithm, illustrates a

---

**Algorithm 2** Distributed Coverage Control with Field Estimation

---

Given information:  $\hat{\theta}(0), Q(0), R(0), S(0)$ .

Return:  $p$ .

If  $k = 0$ , execute

*Initialisation:*

- 1: Initialise the estimator  $\hat{x}$  and covariance  $P$ :

$$\begin{aligned}\tilde{\theta} &= [\hat{\theta}(0)^\top \ 0_p \ 0_q \ 0_r]^\top, \text{ for } \hat{\theta}(0) = \mathbb{E}[\theta(0)], \\ \tilde{P} &= \text{diag}(P(0), Q(0), R(0), S(0)), \text{ for } P(0) = \mathbb{E}[(\theta(0) - \hat{\theta}(0))(\theta(0) - \hat{\theta}(0))^\top].\end{aligned}$$

- 2: Calculate Voronoi partition  $\mathcal{R}$ .

If  $k = 1, 2, \dots, T_k, T_k < \infty$ , execute

*Control Update:*

- 1: Calculate the new centroid using Eq. (123).
- 2: Calculate the control input using Eq. (120).
- 3: Update the position of robot

$$\dot{p} = g(p, u).$$

- 4: Update the Voronoi partition  $\mathcal{R}$ .

*Estimation update:*

- 1: With  $\hat{x} = \hat{\theta}$ , execute estimation procedures in Algorithm 1.
  - 2: Assign  $\hat{\theta} = \hat{x}$ .
- 

drawback arising from using the modified consensus observer in [25]. The second scenario, referred to as the distributed Kalman filter algorithm, demonstrates the performance of the proposed distributed nonlinear Kalman filter given in Algorithm 2. The simulations are performed using Python programming language on a computer with a Linux-based operating system, 2.5-GHz processor, and 4-GB RAM.

Given  $n = 12$  robots randomly scattered in a square area  $\mathcal{Q}$ . We consider the continuous-time dynamics of mobile robots can be represented as a single-integrator system, i.e.,

$$\dot{p} = u, \tag{132}$$

where  $p = [x \ y]^T$  denotes the  $x - y$  position, while  $u = [u_x \ u_y]^T$  is the  $x - y$

control input. The boundaries of the area are  $\{(0, 0), (0, 1), (1, 1), (1, 0)\}$ ; and the information distribution is in the form of

$$\phi(x, y) = \mathcal{K}(q)^\top \theta. \quad (133)$$

Following the parametrisation in (121), the working area is divided into  $m = 4$  partitions; and the kernel function,  $\mathcal{K}(q) = [\mathcal{K}_1, \mathcal{K}_2, \dots, \mathcal{K}_m]$ , has elements given by

$$\mathcal{K}_l = \frac{\exp(-\frac{1}{2}(q - \mu_l)^\top \Sigma_l^{-1}(q - \mu_l))}{\sqrt{(2\pi)^d |\Sigma_l|}},$$

where  $d$  is the dimension of the environment - in this case  $d = 2$ . The elements of the vectorised kernel functions have peaks whose values are given by  $\mu_1 = [0.25, 0.25]^\top$ ,  $\mu_2 = [0.25, 0.75]^\top$ ,  $\mu_3 = [0.75, 0.25]^\top$ , and  $\mu_4 = [0.75, 0.75]^\top$ . The variances of all partitions are similar with  $\Sigma_l = 0.02I_2$ , for  $l = \{1, 2, 3, 4\}$ , where  $I_2 \in \mathbb{R}^{2 \times 2}$  is an identity matrix. The target value of the weights to track is  $\theta = [120.0, 2.5, 2.5, 160.0]^\top$ . Since the mobile robot and the modified-consensus protocol are originally in a continuous-time system, in this work we choose to perform the simulation with sampling time  $0.1s$ .

The parameters required by the coverage controller and the initial values of the estimated weights,  $\hat{\theta}_i$ , in these two simulations are made identical. Since the modified-consensus observer requires some parameters to be given, we choose  $\gamma = 3.63$ ,  $\zeta = 0.6$  and  $k = 1.5$  to show the oscillation caused by large gains. The gains chosen here In the distributed Kalman filter, no static constants are set other than the initial values of estimator and covariance matrix.

By applying the modified-consensus observer algorithm in the first scenario and the distributed Kalman filter in the second one, we obtain the trajectory of robots and the Voronoi partitions illustrated in Figs. 1a and 1b, respectively. The initial positions, final positions and estimated centroids are indicated using red 'x', blue 'o' and orange 'x'. From Fig. 1a, it can be seen that there is an agent that do not converge to the centroid; while according to Fig. 1b all robots can successfully converge to the estimated centroids.

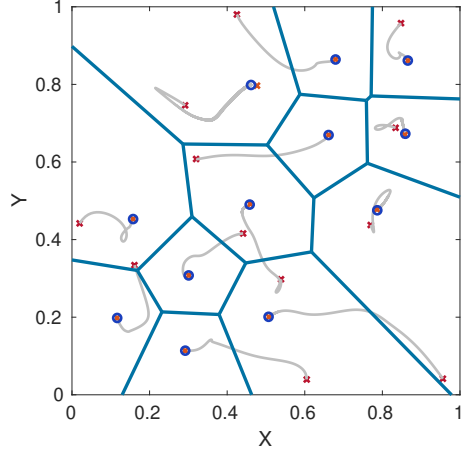
Related to the performance of field estimation algorithm, the implementation of the modified-consensus observer yields the convergence results of the estimated weights  $\hat{\theta}_i$  presented in Fig. 2a; while the distributed Kalman filter yields the convergence results of the estimated weights given in Fig. 2b. These figures shows that the distributed Kalman filter has successfully driven the estimated weights

to some values close to the target weights. As comparison, the estimation using modified-consensus observer shows high-amplitude oscillation with constant mean of several estimated weights.

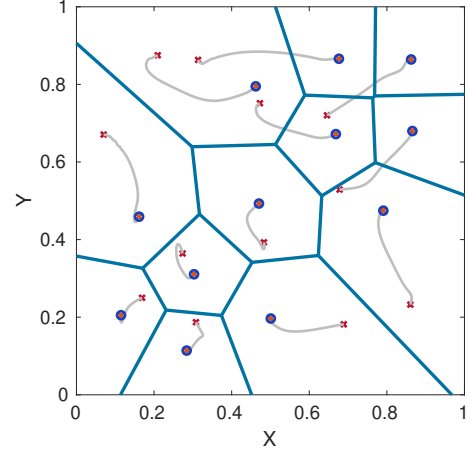
The performance of coverage control using the modified-consensus observer as its field estimator is illustrated in Fig. 3a for convergence of the norm of the tracking errors  $\|p_i - \hat{\mathcal{C}}_{\mathcal{R}_i}\|$ , and Fig. 4a for the convergence of the objective function; while results of the coverage control using the distributed Kalman filter are depicted in Fig. 3b for the convergence of the norm of the errors and Fig. 4b for the convergence of the objective functions. Since the weights obtained using the modified-consensus observer do not converge to zero, the errors also do not converge to zero. According to Figures 3a, the estimated objective function using the modified-consensus observer has steady state error with the true objective function. As comparison, using the distributed Kalman filter, the systems are able to minimise the errors and quickly track the true objective function.

Empirically, in addition to the simulation results presented above, we also calculate the average computation time of 500 iterations of these presented scenarios. Although there exist variations in the computation time for each iteration, the average time required to perform the iterations of coverage control with distributed Kalman filter always less than the one with modified-consensus observer. The average of the computation time of the first scenario with the modified-consensus observer varies along the process with average of 37.0349s per iteration, while the second scenario with our algorithm takes 30.5859s per iteration in average.

From the illustrated figures, it can be concluded that: 1) Both estimators can be used to estimate the information distribution; 2) It is true that the simulation result in [25] has shown non-oscillating results with some chosen gains. However, the algorithm still requires the gains to be tuned but no exact upper boundary of the estimation gains has been given. Failure to pick the right gains may lead to oscillating performance and steady-state tracking error of the objective function as shown in this comparative simulation; 3) Unlike the algorithm in [25], the proposed distributed nonlinear Kalman filter needs no constant gains to be tuned to produce stable performance; 4) The computation time, which can also indicate the computation burden, of the distributed Kalman filter is faster than the modified-consensus observer algorithm in [25]. This might be caused by a number of surface integrations performed to adjust some variables in the modified-consensus observer where the distributed Kalman filter requires no additional surface integrations in the estimation process.

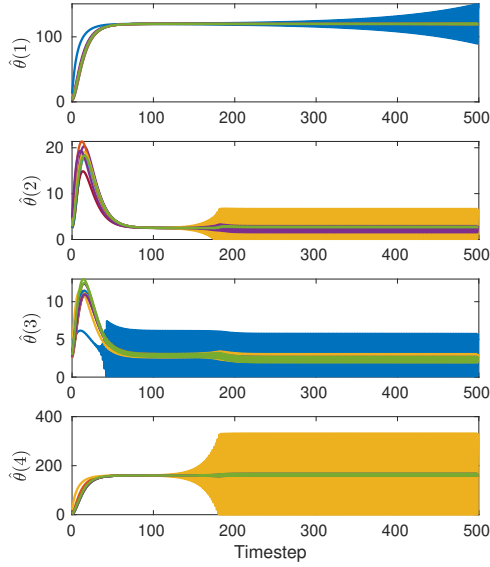


(a) Using modified-consensus observer.

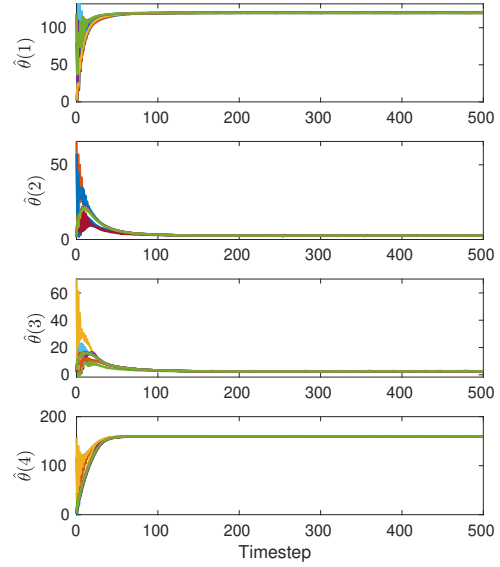


(b) Using Kalman-consensus estimator.

Figure 1: Position trajectories and optimal centroidal Voronoi regions.



(a) Using modified-consensus observer.



(b) Using Kalman-consensus estimator.

Figure 2: Convergence result of the estimated weights of the density function  $\hat{\theta}_i, \forall i \in \mathcal{V}_n$

## 7. Conclusions

In this paper, the distributed nonlinear Kalman filter with general communication scheme has been presented to estimate a dynamical process with additive



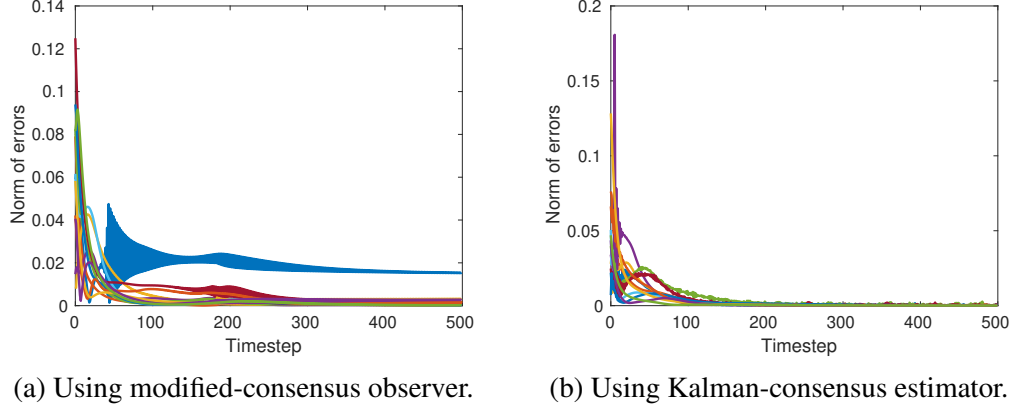


Figure 3: Convergence result of error  $\|p_i - \hat{C}_{\mathcal{R}_i}\|, \forall i \in \mathcal{V}_n$ .

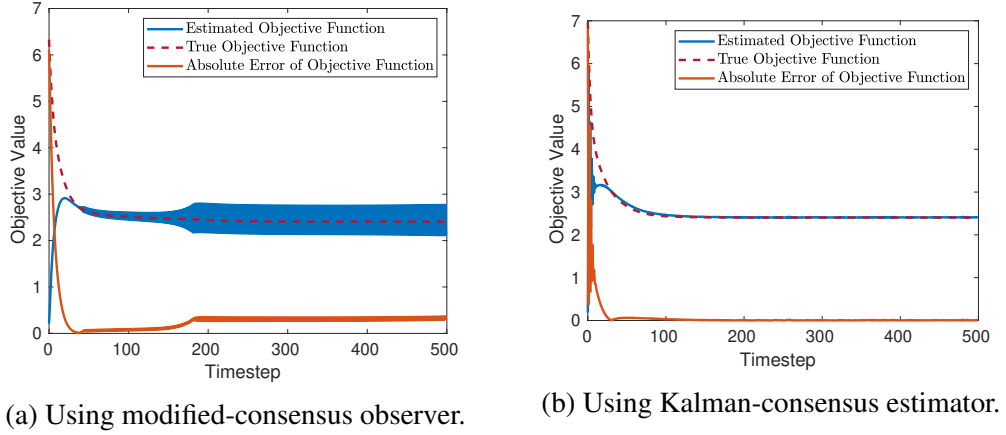


Figure 4: Convergence result of the objective function.

white Gaussian noises in the system, measurement and communication. The optimal Kalman and communication gains have been provided such that the estimator has capability of using measurement and communicated information to produce an estimate value. Using this mechanism, the consensus protocol combined with Kalman filter, which was reported in literature, could be considered as a special case of the proposed algorithm. After designing the estimation algorithm, we have analytically demonstrated that the estimate error is exponentially bounded in mean square with regards to some boundaries. As an example, the algorithm has also been applied to coverage problem scenarios with a previously-unidentified information distribution. Combined with the coverage control protocol in our

previous results [22], the proposed algorithm was used to estimate the density function and find the optimal deployment of robots. Two scenarios of numerical experiments have been carried out as a comparison with the existing method used for field estimation in coverage control problem. The results have shown that the proposed distributed nonlinear Kalman filter algorithm have outperformed the modified-consensus observer in [25] by successfully estimating the unknown density function of an environment indicated, driving all agents to the optimal centroid positions and minimise the true objective function without oscillations. Moreover, the simulation results indicated that the proposed algorithm needs less computation time than the existing modified-consensus observer.

## References

- [1] R. Mehra, On the identification of variances and adaptive kalman filtering, *IEEE Transactions on Automatic Control* 15 (2) (1970) 175–184 (April 1970).
- [2] K. Reif, S. Gunther, E. Yaz, R. Unbehauen, Stochastic stability of the discrete-time extended kalman filter, *IEEE Transactions on Automatic Control* 44 (4) (1999) 714–728 (Apr 1999).
- [3] E. A. Wan, R. V. D. Merwe, The unscented kalman filter for nonlinear estimation, in: *Proceedings of the IEEE 2000 Adaptive Systems for Signal Processing, Communications, and Control Symposium (Cat. No.00EX373)*, 2000, pp. 153–158 (2000).
- [4] S. J. Julier, J. K. Uhlmann, Unscented filtering and nonlinear estimation, *Proceedings of the IEEE* 92 (3) (2004) 401–422 (Mar 2004).
- [5] I. Arasaratnam, S. Haykin, T. R. Hurd, Cubature kalman filtering for continuous-discrete systems: Theory and simulations, *IEEE Transactions on Signal Processing* 58 (10) (2010) 4977–4993 (Oct 2010).
- [6] L. Li, Y. Xia, [Ukf-based nonlinear filtering over sensor networks with wireless fading channel](#), *Information Sciences* 316 (2015) 132 – 147, *nature-Inspired Algorithms for Large Scale Global Optimization* (2015).
- [7] M. Oussalah, J. D. Schutter, [Possibilistic kalman filtering for radar 2d tracking](#), *Information Sciences* 130 (1) (2000) 85 – 107 (2000).

- [8] G. Hao, S. li Sun, Y. Li, [Nonlinear weighted measurement fusion unscented kalman filter with asymptotic optimality](#), Information Sciences 299 (2015) 85 – 98 (2015).
- [9] R. Olfati-Saber, Distributed kalman filter with embedded consensus filters, in: Proceedings of the 44th IEEE Conference on Decision and Control, 2005, pp. 8179–8184 (Dec 2005).
- [10] R. Olfati-Saber, Kalman-consensus filter : Optimality, stability, and performance, in: Proceedings of the 48h IEEE Conference on Decision and Control (CDC) held jointly with 2009 28th Chinese Control Conference, 2009, pp. 7036–7042 (Dec 2009).
- [11] F. S. Cattivelli, A. H. Sayed, Diffusion strategies for distributed kalman filtering and smoothing, IEEE Transactions on Automatic Control 55 (9) (2010) 2069–2084 (Sept 2010).
- [12] G. Battistelli, L. Chisci, [Stability of consensus extended kalman filter for distributed state estimation](#), Automatica 68 (Supplement C) (2016) 169 – 178 (2016).
- [13] S. Wang, W. Ren, Fully distributed nonlinear state estimation using sensor networks, in: 2017 American Control Conference (ACC), 2017, pp. 2562–2567 (May 2017).
- [14] Q. Tan, X. Dong, Q. Li, Z. Ren, Weighted average consensus-based cubature kalman filtering for mobile sensor networks with switching topologies, in: 2017 13th IEEE International Conference on Control Automation (ICCA), 2017, pp. 271–276 (July 2017).
- [15] W. Li, G. Wei, F. Han, Y. Liu, Weighted average consensus-based unscented kalman filtering, IEEE Transactions on Cybernetics 46 (2) (2016) 558–567 (Feb 2016).
- [16] Y. Niu, L. Sheng, Distributed consensus-based unscented kalman filtering with missing measurements, in: 2017 36th Chinese Control Conference (CCC), 2017, pp. 8993–8998 (July 2017).
- [17] A. Okabe, A. Suzuki, [Locational optimization problems solved through Voronoi diagrams](#), European Journal of Operational Research 98 (3) (1997) 445–456 (May 1997).

- [18] A. Okabe, Spatial tessellations: concepts and applications of Voronoi diagrams, 2nd Edition, Wiley series in probability and statistics, Wiley, Chichester ; New York, 2000 (2000).
- [19] M. Pavone, A. Arsie, E. Frazzoli, F. Bullo, Distributed algorithms for environment partitioning in mobile robotic networks, *IEEE Transactions on Automatic Control* 56 (8) (2011) 1834–1848 (Aug 2011).
- [20] J. Cortes, S. Martinez, T. Karatas, F. Bullo, [Coverage control for mobile sensing networks](#), *IEEE Transactions on Robotics and Automation* 20 (2) (2004) 243–255 (Apr. 2004).
- [21] S. G. Lee, Y. Diaz-Mercado, M. Egerstedt, [Multirobot control using time-varying density functions](#), *IEEE Transactions on Robotics* 31 (2) (2015) 489–493 (Apr. 2015).
- [22] H. Tnunay, Z. Li, C. Wang, Z. Ding, Distributed collision-free coverage control of mobile robots with consensus-based approach, in: 2017 13th IEEE International Conference on Control Automation (ICCA), 2017, pp. 678–683 (July 2017).
- [23] M. Schwager, F. Bullo, D. Skelly, D. Rus, [A ladybug exploration strategy for distributed adaptive coverage control](#), *IEEE*, 2008, pp. 2346–2353 (May 2008).
- [24] M. Schwager, J.-J. Slotine, D. Rus, [Consensus learning for distributed coverage control](#), *IEEE*, 2008, pp. 1042–1048 (May 2008).
- [25] M. Schwager, M. P. Vitus, S. Powers, D. Rus, C. J. Tomlin, Robust adaptive coverage control for robotic sensor networks, *IEEE Transactions on Control of Network Systems* 4 (3) (2017) 462–476 (Sept 2017).
- [26] R. A. Horn, C. R. Johnson (Eds.), *Matrix Analysis*, Cambridge University Press, New York, NY, USA, 1986 (1986).
- [27] T.-J. Tarn, Y. Rasis, Observers for nonlinear stochastic systems, *IEEE Transactions on Automatic Control* 21 (4) (1976) 441–448 (Aug 1976).
- [28] E. A. Wan, R. V. D. Merwe, *The Unscented Kalman Filter, Kalman Filtering and Neural Networks* (2001) 62 pages (2001).

- [29] K. Xiong, H. Zhang, C. Chan, [Performance evaluation of ukf-based nonlinear filtering](#), Automatica 42 (2) (2006) 261 – 270 (2006).
- [30] M. Geist, O. Pietquin, [Kalman Temporal Differences](#), Journal of Artificial Intelligence Research 39 (2010) 483–532 (Oct. 2010).
- [31] S. P. Boyd, L. Vandenberghe, Convex optimization, Cambridge University Press, Cambridge, UK; New York, 2004 (2004).
- [32] R. Sanner, J.-J. Slotine, [Gaussian networks for direct adaptive control](#), IEEE Transactions on Neural Networks 3 (6) (1992) 837–863 (Nov. 1992).