Distributed Optimal Consensus Filter for Target Tracking in Heterogeneous Sensor Networks

Shanying Zhu, Student Member, IEEE, Cailian Chen, Member, IEEE, Wenshuang Li, Student Member, IEEE, Bo Yang, Member, IEEE, and Xinping Guan, Senior Member, IEEE

Abstract—This paper is concerned with the problem of filter design for target tracking over sensor networks. Different from most existing works on sensor networks, we consider the heterogeneous sensor networks with two types of sensors different on processing abilities (denoted as type-I and type-II sensors, respectively). However, questions of how to deal with the heterogeneity of sensors and how to design a filter for target tracking over such kind of networks largely remain unexplored. We propose in this paper a novel distributed consensus filter to solve the tracking problem. Two criteria, namely, unbiasedness and optimality, are imposed for the filter design. The so-called sequential design scheme is then presented to tackle the heterogeneity of sensors. The minimum principle of Pontryagin is adopted for type-I sensors to optimize the estimation errors. As for type-II sensors, the Lagrange multiplier method coupled with the generalized inverse of matrices is then used for filter optimization. Furthermore, it is proven that convergence property is guaranteed for the proposed consensus filter in the presence of process and measurement noise. Simulation results have validated the performance of the proposed filter. It is also demonstrated that the heterogeneous sensor networks with the proposed filter outperform the homogenous counterparts in light of reduction in the network cost, with slight degradation of estimation performance.

Index Terms—Heterogeneous sensor network, optimal consensus filter, target tracking, unbiased estimate.

I. INTRODUCTION

To locate and track a moving target is crucial for many applications such as robotics, surveillance, monitoring, and security for large-scale complex environments [1]–[3]. In such scenarios, a number of sensors can be employed in order to improve the tracking accuracy and increase the size of the surveillance area in a cooperative manner. Basically, these sensors have modest capabilities of sensing, computation, and multihop wireless communication. Equipped with these capabilities, the sensors can self-organize to form a network that is capable of sensing and processing spatial and temporal dense data in the monitored area.

However, due to radio wave propagation loss in wireless channels, reliable communications between sensors can only be ensured within a short distance. As a result, one of the main challenges for target tracking over sensor networks is that only local information from neighboring sensors is available at each sensor. In this case, the traditional centralized methods (e.g., Kalman filter and extended Kalman filter [1]) for target tracking are not applicable. One possible way is to decentralize the task over the whole network, which brings forward distributed algorithms. The major advantages of distributed algorithms are that: 1) they can reduce the cost of transmitting all data to the fusion center, and 2) they exhibit resilience against sensor failures, thus achieving robustness of the network. A decentralized version of a Kalman filter was proposed in [4], [5], whose main idea is to decentralize the Kalman filtering task to each sensor. However, this algorithm is not scalable with the network size since each sensor needs to communicate with all others to compute its local estimate. A more general setting of the distributed Kalman filter was presented in [5], where a proximity graph is considered, i.e., each sensor only needs to share its information with neighboring sensors. Additionally, a diffusion step is added after the estimate update step to improve the performance of the distributed Kalman filter. Recently, consensus algorithms of multiagent systems have been proven to be effective tools for performing network-wide distributed tasks [6], [7]. Motivated by these algorithms, Olfati-Saber introduced a scalable distributed Kalman-consensus filter 68 (KCF) [8]. It is modified from the Kalman filter by inserting a consensus term to reduce the disagreements of estimates among the sensors. However, it is far from optimal with respect to the 71 error covariance, which might cause unacceptable estimation errors. The optimality and stability analysis in its discrete-time form was further studied in [9]. An alternative distributed 74 Kalman filter based on consensus algorithms was proposed in [10], which is composed of two stages, i.e., a Kalman-like measurement update and an estimate fusion using a consensus matrix. However, only a scalar system was considered, and the optimization of a Kalman gain and a consensus matrix is, in general, nonconvex. In [11], a pinning observer was designed to solve the filtering problem in the case that sensors can only observe partial target information. The distributed filtering with partial information between sensors was investigated in [12].
An optimization approach was also used to solve distributed estimation problems. For example, [13] reformulates the centralized Kalman filter and smoother for distributed operation through the alternating direction method of multipliers. The proposed algorithm offers any-time optimal state estimates based only on local information.

It is noted that most of the aforementioned works deal with homogeneous sensor networks, i.e., all sensors possess identical communication and computation capability, and so on. It is shown that the homogeneous architecture suffers from poor fundamental limits and performance [14]. On the other hand, heterogeneous networks have been experimentally shown to be superior to homogeneous ones due to their potential to increase the network lifetime and reliability [15].

Taking these into considerations, we deal with the target tracking problem in a heterogeneous sensor network framework. Two types of sensors different on processing abilities, which are denoted as type-I and type-II sensors, are present in the network. Type-I sensors are of more computational ability than type-II ones. Questions of how to deal with the heterogeneity of sensors to facilitate the filter design over such kind of sensor networks remain largely unexplored. To the best of the authors’ knowledge, there is no filter design method reported in the open literature for target tracking in such kind of heterogeneous sensor networks. Our main objective is to propose a distributed optimal consensus filter (DOCF) for target tracking relying only on neighbor-to-neighbor communication among sensors. However, the heterogeneity of sensors takes new challenges for the optimal filter design: 1) The coexistence of two types of sensors in the network needs filter forms for different types of sensors. This makes it impossible to design the optimal filter in a unified framework as in [9], [10], and [13] with identical expressions; 2) The convergence is much more difficult to analyze due to the quite different stochastic representations of the target model and the heterogeneous filter models. The analysis is even untractable without delicate design of the filter. The main contributions of this paper are twofold.

- We propose a distributed consensus filter for target tracking in heterogeneous sensor networks. The filter parameters are designed according to the performance requirements on unbiasedness and optimality. A sufficient and necessary condition to guarantee the unbiasedness is established. Unlike the homogeneous scenario considered in [9] and [13], we develop the so-called sequential design approach to achieve the optimality with sensor heterogeneity. Specifically, for type-I sensors, the optimization problem is first casted into an optimal control problem. Then, the minimum principle of Pontryagin is applied to solve this problem. Afterward, for type-II sensors, the Lagrange multipler and least square method is used to obtain the optimal weights. Moreover, the discrete-time DOCF is proposed by discretization and approximation.

- The convergence property of the proposed filter is investigated. We analyze the estimation error dynamics via subtly rewriting the system in the Itô stochastic framework. By using the stochastic Lyapunov method and the stopping time technique, we show that the estimation error of the consensus filter is exponentially bounded in mean square. Simulations have validated the theoretical analysis.

The remainder of this paper is organized as follows: Section II presents the network model of heterogeneous sensor networks, the target model, and the distributed consensus filter. In Section III, we determine the structure of the distributed consensus filter from two aspects, namely, unbiasedness and optimality of the estimates. The mean-square performance of the consensus filter is provided in Section IV. In Section V, simulation results are given to illustrate the effectiveness of the proposed consensus filter with some comparison with some other existing filters. Finally, Section VI concludes this paper.

II. PROBLEM FORMULATION

A. Target and Network Models

The system considered for target tracking problem in this paper is composed of \( N \) sensors and a moving target in a 157 monitored field. The purpose of the sensors is to cooperatively trace the behavior of the target based only on neighbor-to-neighbor communication.

Consider a moving target with the following dynamics:

\[
x(t) = Ax(t) + Bu(t)
\]

where \( A \in \mathbb{R}^{n \times n} \) and \( B \in \mathbb{R}^{n \times m_1} \) are constant matrices, \( x(t) \in \mathbb{R}^n \) is the state vector, and \( u(t) \in \mathbb{R}^{m_1} \) is the ambient \( 163 \) noise, which is assumed to be zero-mean and white. The initial state \( x_0 \) is a Gaussian random variable with known initial mean \( \mathbb{E}\{x_0\} = \bar{x}_0 \) and positive definite covariance \( \mathbb{E}\{(x_0 - 166 \bar{x}_0)(x_0 - \bar{x}_0)^T\} = \Pi_0 > 0 \).

Two types of sensors with different computational capabilities are deployed in the monitored field to locate and track the moving target. Type-I sensors have high processing abilities, whereas type-II sensors are low-end ones. Furthermore, we assume that only type-I sensors are equipped with onboard 172 sensing units. They can sense the environment and observe the 173 target with noisy measurements of the target. Thus

\[
y_i(t) = C_i x(t) + v_i(t) \quad \forall i \in \mathcal{I}
\]

where \( C_i \in \mathbb{R}^{n \times m_2} \) is the observation matrix, and \( v_i(t) \in \mathbb{R}^{m_2} \) is the zero-mean and white measurement noise. As for type-II 176 sensors, they cannot measure the state of the target, which means that all the information about the target at each type-II sensor is directly or indirectly obtained from the type-I sensors. 179

The sensors in the network are coupled by an undirected 180 communication topology \( \mathcal{G} = (\mathcal{V}, \mathcal{E}) \), where \( \mathcal{V} = \mathcal{I} \cup \mathcal{I}^c = 181 \{1, 2, \ldots, N\} \) is the set of sensors; \( \mathcal{I} \) and \( \mathcal{I}^c \) are the set of type-I 182 and type-II sensors, respectively; and \( \mathcal{E} \subset \mathcal{V} \times \mathcal{V} \) is the set of 183 communication links, over which local information about the 184 target is exchanged. Without loss of generality, let the first \( M \) 185 sensors belong to \( \mathcal{I} \) and the rest \( N-M \) sensors belong to \( \mathcal{I}^c \). In this paper, we assume that each sensor transmits at constant power \( p_T \) and the receiver has ambient noise power \( n_R \). Since 188 path loss is inevitably encountered as the radio wave propagates 189 through the environments, the signal power decreases with the 190 distance captured by the path loss exponent. Then, the signal 191
transmitted from sensor $i$ can be successfully received by sensor $j$ only if the signal-to-noise ratio (SNR) satisfies
\[
\text{SNR} \triangleq \frac{Pr}{n_R d_{ij}^\eta} \geq \rho
\]
where $d_{ij}$ is the distance between the sensors $i$ and $j$, $\eta$ is the path loss exponent (typically, $2 \leq \eta \leq 5$), and $\rho > 0$ is the threshold. Thus, reliable wireless communications between sensors can only be ensured within a distance of $r \triangleq \sqrt{\frac{Pr}{n_R \rho}}$. This way, the link $(i, j) \in \mathcal{E}$ exists if and only if $d_{ij} \leq r$. Moreover, the set of neighbors of sensor $i$ can be defined as $\mathcal{N}_i \triangleq \{j \in \mathcal{V} : d_{ij} \leq r\}$ for all $i \in \mathcal{V}$.

201 B. Distributed Consensus Filter for Heterogeneous Networks

As mentioned, the Kalman filter and its distributed generalizations for sensor networks [3], [4] are not applicable to target tracking in the heterogeneous sensor networks. We have proposed in [16] a distributed consensus filter for heterogeneous sensor networks.

In this paper, the method in [16] is developed to a novel DOCF for the target tracking scenario. The distributed consensus filter applicable to heterogeneous networks is given as follows. For type-I sensor $i \in \mathcal{I}$
\[
\dot{x}_i(t) = F_i(t) \dot{x}_i(t) + G_i(t)y_i(t) + H_i(t) \sum_{j \in \mathcal{N}_i} a_{ij} [\dot{x}_j(t) - \dot{x}_i(t)]
\]
and for type-II sensor $i \in \mathcal{I}^c$
\[
\dot{x}_i(t) = \sum_{j \in \mathcal{N}_i} \gamma_{ij}(t) \dot{x}_j(t)
\]
where $\dot{x}_i(t) \in \mathbb{R}^n$ is the estimate of the target state $x(t)$, $F_i \in \mathbb{R}^{n \times n}$ and $G_i \in \mathbb{R}^{n \times m_2}$ are the filter matrices, $H_i \in \mathbb{R}^{n \times n}$ is the consensus gain matrix, $a_{ij} = \sqrt{Pr/d_{ij}^\eta}$ represents the received signal strength at sensor $i$ transmitted from sensor $j$, and $\gamma_{ij} > 0$ are constant weights.

Equations (3) and (4) constitute the distributed consensus filter. In the evolution, the type-I sensors sense and measure the location of the target. Meanwhile, they receive the local estimates from their neighbors. With this information, it is possible for the type-I sensors to update their own estimates via (3). Then, the type-I sensors transmit the local estimates to their neighbors. After receiving the local estimates, the type-II sensors scale all the received data in order to find the weighted average in (4). Note that, in the preceding process, communication only takes place among neighboring sensors; thus, the consensus filter (3) and (4) is totally distributed and thus scalable with the network size. Fig. 1 depicts a schematic of the distributed consensus filter at the sensor level.

Remark 1: From (3) and (4), it can be clearly seen that type-I sensors have more powerful computational ability, and thus, they can perform more complicated tasks, such as filtering of its measurement and the local estimates from its neighbors. While the abilities of type-II sensors are much limited, they fuse the data to form their own estimates by only simply weighting all the incoming data.

As for the distributed consensus filter (3) and (4), there are three sets of parameters to be determined, namely, $F_i$, $G_i$, and $\gamma_{ij}$, $i, j \in \mathcal{V}$. To tackle the heterogeneity of (3) and (4), we propose the sequential design approach to determine these parameters in the next section in light of the unbiasedness and optimality of the estimates.

III. Structure of Consensus Filter: Unbiasedness and Optimality

Here, we address the problem of determination of filter matrices $F_i$, $G_i$, and $\gamma_{ij}$ to meet the requirement of unbiasedness, i.e., $\mathbb{E} \{e_i(t)\} = 0$, $\forall i \in \mathcal{V}$, $t \geq 0$, where $e_i(t)$ is the estimation error
\[
e_i(t) \triangleq \hat{x}_i(t) - x(t).
\]

Lemma 1: For each $i \in \mathcal{V}$, let the initial estimate $\hat{x}_i(0) = \hat{x}_0$, then $\hat{x}_i(t)$, $\forall t > 0$, is unbiased if and only if $d_{ij} \in \mathcal{I}$ for type-I sensor, $F_i(t) = A - G_i(t)C_i$ and $2$ for type-II sensor, $\sum_{j \in \mathcal{N}_i} \gamma_{ij}(t) = 1$.

Proof: See Appendix A. 

A. Unbiased Estimate

The requirement of unbiasedness is widely considered for the purpose of theoretical derivation, but in some engineering applications, it has been shown to be desirable from a practical viewpoint [18], [19].

The next lemma presents a way to choose the parameters $F_i$, $G_i$, and $\gamma_{ij}$ to meet the requirement of unbiasedness, i.e., $\mathbb{E} \{e_i(t)\} = 0$, $\forall i \in \mathcal{V}$, $t \geq 0$, where $e_i(t)$ is the estimation error
\[
e_i(t) \triangleq \hat{x}_i(t) - x_i(t).
\]

Lemma 1: For each $i \in \mathcal{V}$, let the initial estimate $\hat{x}_i(0) = \hat{x}_0$, then $\hat{x}_i(t)$, $\forall t > 0$, is unbiased if and only if $d_{ij} \in \mathcal{I}$ for type-I sensor, $F_i(t) = A - G_i(t)C_i$ and $2$ for type-II sensor, $\sum_{j \in \mathcal{N}_i} \gamma_{ij}(t) = 1$.

Proof: See Appendix A. 

As a result, it follows from Lemma 1 that the unbiased filter for type-I sensor $i \in \mathcal{I}$ is given by
\[
\dot{x}_i(t) = (A - G_i(t)C_i) \dot{x}_i(t) + G_i(t)y_i(t)
\]
\[+ H_i(t) \sum_{j \in \mathcal{N}_i} a_{ij} [\dot{x}_j(t) - \dot{x}_i(t)].
\]
As for the type-II sensor, from Lemma 1 and Proposition 1 [16], the unbiased filter (4) has the following form:

$$\hat{x}_i(t) = \sum_{j \in \tilde{I}} \gamma_{ij}(t) \hat{x}_j(t), \quad i \in \tilde{I}$$

(6)

To facilitate the filter design, we give some standard assumptions.

$$\left\{\begin{array}{l}
\sum_{j \in \tilde{I}} \gamma_{ij}(t) > 0, \quad k = 1, \ldots, p_i, \\
\sum_{j \in \tilde{I}} \gamma_{ij}(t) = 1, \quad k = p_i + 1, \ldots, |\tilde{I}|,
\end{array}\right.$$  

(7)

2.2 B. Optimization of the Gain Matrix $G_i$ and Weights $\tilde{\gamma}_{ij}$

Optimal filters possess some important merits, such as being robust in their maintenance of performance standards [18]. In the target tracking scenario, optimal filters for homogeneous sensor networks were considered in [9] and [13]. In our setting, however, the heterogeneity of sensors presents a challenge for the optimal filter design. We propose the sequential design approach to tackle this difficulty.

To facilitate the filter design, we give some standard assumptions. The random variables $x_0$, $w(t)$, and $v_i(t)$, $i \in \tilde{I}$, are 282 independent, and

$$\mathbb{E}\{w(t)w^T(\tau)\} = Q(t)\delta(t - \tau)$$

$$\mathbb{E}\{v_i(t)v_j^T(\tau)\} = R_{ij}(t)\delta(t - \tau) \quad \forall i, j \in \tilde{I}$$

(8)

(9)

where $\delta(\cdot)$ is the Dirac delta function. Moreover, assume that $R_i(t) \triangleq R_{ii}(t), \forall i \in \tilde{I}$, is positive definite.

Define the error covariance matrix as $P_{ij}(t) \triangleq \mathbb{E}\{e_i(t)e_j^T(t)\}$, for each pair $(e_i(t), e_j(t))$, $i, j \in \mathcal{V}$, and denote $P_{Ni,j}(t) \triangleq \sum_{r \in N_i} a_{ir}[P_{ij}(t) - P_{ij}(t)]$ and $P_{N_i}(t) \triangleq \sum_{j \in \mathcal{V}} a_{ij}[P_{ij}(t) - P_{ij}(t)]$. Then, a standard argument shows that $P_{ij}(t)$ satisfies (see Appendix B) the cases here.

290 **Case i** $i \in \tilde{I}$ and $j \in \tilde{I}$.

$$\hat{P}_{ij}(t) = (A - G_i(t)C_i) P_{ij}(t) + P_{ij}(t) (A - G_j(t)C_j)^T$$

$$+ H_i(t) P_{Ni,j}(t) + P_{Ni,j}(t) H_j^T(t)$$

$$+ G_i(t) R_{ij}(t) G_j^T(t) + BQ(t)B^T.$$  

(10)

291 **Case ii** $i \in \tilde{I}$ and $j \in \mathcal{I}^c$.

$$\hat{P}_{ij}(t) = (A - G_i(t)C_i) P_{ij}(t)$$

$$+ \sum_{k \in \mathcal{I}} \tilde{\gamma}_{ik}(t) P_{jk}(t) (A - G_k(t)C_k)^T$$

$$+ H_i(t) P_{Ni,j}(t) + \sum_{k \in \mathcal{I}} \tilde{\gamma}_{ik}(t) P_{iN_k}(t) H_k^T(t)$$

$$+ \sum_{k \in \mathcal{I}} \tilde{\gamma}_{ik}(t) G_i(t) R_{ik}(t) G_k^T(t) + BQ(t)B^T.$$  

(11)

292 **Case iii** $i \in \mathcal{I}^c$ and $j \in \tilde{I}$.

$$\hat{P}_{ij}(t) = P_{ij}(t) (A - G_j(t)C_j)^T$$

$$+ \sum_{k \in \mathcal{I}} \tilde{\gamma}_{ik}(t) (A - G_k(t)C_k) P_{kj}(t)$$

$$+ P_{iN_k}(t) H_j^T(t) + \sum_{k \in \mathcal{I}} \tilde{\gamma}_{ik}(t) H_k(t) P_{N_k,j}(t)$$

$$+ \sum_{k \in \mathcal{I}} \tilde{\gamma}_{ik}(t) G_k(t) R_{kj}(t) G_j^T(t) + BQ(t)B^T.$$  

(12)

293 **Case iv** $i \in \mathcal{I}^c$ and $j \in \mathcal{I}^c$.

$$P_{ij}(t) = \sum_{k \in \mathcal{I}^c} \gamma_{ik}(t) P_{kj}(t).$$  

(13)

From a physical point of view, it is desirable to have an unbiased estimate with minimum covariance $P_i \triangleq P_{ii}$. In the following, however, we consider the perturbed covariance matrix $\hat{P}_i$ in view of convergence performance:

$$\hat{P}_i(t) = (A - G_i(t)C_i) \hat{P}_i(t) + \hat{P}_i(t) (A - G_i(t)C_i)^T$$

$$+ H_i(t) \hat{P}_{Ni,i}(t) + \hat{P}_{Ni,i}(t) H_i^T(t)$$

$$+ G_i(t) R_i(t) G_i^T(t) + BQ(t)B^T + W_i.$$  

(14)

where $W_i$ is positive definite, $\hat{P}_{Ni,j}(t) \triangleq \sum_{r \in N_i} a_{ir} [\hat{P}_{ij}(t) - 298 P_{ij}(t) \hat{P}_{ij}(t)]$, $\hat{P}_{Ni}(t) \triangleq \sum_{j \in \mathcal{V}} a_{ij} [\hat{P}_{ij}(t) - P_{ij}(t)]$, and $\hat{P}_{ij}$ for all 299 $i, j \in \mathcal{V}$ are likewise determined as $P_{ij}$ in four cases. Denote 300 the corresponding equations as (11’), (11’’), (12’), and (13’), respectively. Moreover, we set $\hat{P}_i(0) = P_{ii}(0) = P_0$. 302

We introduce the cost function

$$\mathcal{J}(P_i) = \text{tr} \left[ \hat{P}_i(t) \right], \quad i \in \mathcal{V}$$  

(15)

as the measure of the performance of the filter, where $\text{tr}[\cdot]$ denotes the trace of a matrix. The smaller the value of $\mathcal{J}_i$, the better the consensus filter (5) and (6).

It is observed from (10)–(13) that, with the cost function $\mathcal{J}_i$, matrix $G_i$ can be independently determined from weights $\tilde{\gamma}_{ij}$. This observation makes it possible to sequentially design the parameters $G_i$ and $\tilde{\gamma}_{ij}$ for type-I and type-II sensors, respectively. We first focus on the selection of $G_i$ for the type-I sensors.

**Lemma 2:** For type-I sensor $i \in \mathcal{I}$, the optimal matrix $G_i^* 313$ with respect to the cost function $\mathcal{J}_i(P_i)$ is given by $G_i^*(t) = 314 \hat{P}_i(t)C_i^T R_i^{-1}C_i \hat{P}_i(t)$, where $\hat{P}_i(t)$ is the solution of the matrix equation

$$\hat{P}_i(t) = A \hat{P}_i(t) A^T - \hat{P}_i(t) C_i^T R_i^{-1}(t) C_i \hat{P}_i(t)$$

$$+ H_i(t) \hat{P}_{Ni,i}(t) + \hat{P}_{Ni,i}(t) H_i^T(t)$$

$$+ BQ(t)B^T + W_i.$$  

(16)
TABLE I
PARAMETERS FOR THE UNBIASED DISTRIBUTED CONSENSUS FILTER (3) AND (4)

<table>
<thead>
<tr>
<th>Unbiasedness</th>
<th>type-I sensor</th>
<th>type-II sensor</th>
</tr>
</thead>
<tbody>
<tr>
<td>$F_i(t) = A - G_i(t)C_i$</td>
<td>$\gamma_{ij} = 1$</td>
<td>$\gamma_{ij} = 1$</td>
</tr>
<tr>
<td>Optimality</td>
<td>$G_i^2(t) = P_i^E(t)C_i^T R_i^{-1}(t)$</td>
<td>$\sum_{j \in \mathcal{I}} \gamma_{ij} = 1$</td>
</tr>
</tbody>
</table>

\[ z^*_{ij} = \begin{cases} 1 - \sum_{k=2}^{p_i} z_{jk}, & j = j_1, \\
\cdot & j = j_2, \ldots, j_{p_i}, \\
0, & j = j_{p_i+1}, \ldots, J_{ji} \end{cases} \]  

(17)

where $z^* = [z_{j_1}, z_{j_2}, \ldots, z_{j_{p_i}}]$. The linear equation $U_i z = V_i$, in which

\[ U_i \triangleq \begin{bmatrix} U_{j_2,j_1} & \cdots & U_{j_{p_i},j_1} \\
\cdot & \cdots & \cdot \\
U_{j_{p_i},j_{p_i}} & \cdots & U_{j_{p_i},j_{p_i}} \end{bmatrix}, \quad V_i \triangleq \begin{bmatrix} V_{j_2} \\
\cdot \\
V_{j_{p_i}} \end{bmatrix} \]

and $U_{j_2,j_1} = \text{tr}[\hat{P}_{j_2,j_1} - \hat{P}_{j_2,j_1} - \hat{P}_{j_2,j_1}]$ and $V_{j_2} = \text{tr}[\hat{P}_{j_2,j_1}]$.

Proof: See Appendix D.

In general, there may exist no positive solution or more than one positive solution of $U_i z = V_i$. In this case, we turn to the least squares solution by minimizing the norm of the residual, i.e.,

\[ \min_{z \in \mathbb{R}^p} \|U_i z - V_i\|, \quad \text{s.t.} \quad z > 0. \]  

(18)

Several numerical methods can be used here to solve this problem efficiently such as the active set algorithm [20].

Theorem 1: Consider a heterogeneous sensor network with communication topology $G$ tracking the target with dynamics (1) and (2). Assume that only type-I sensors can observe the target with measurement (2), then the unbiased distributed consensus filter (3) and (4) is optimal with respect to the cost function $J_r(\hat{P}_i)$ and (15) with the parameters given in Table I.

C. Discrete-Time Version of the Consensus Filter: DOCF

In the previous subsections, we have demonstrated how to select the parameters $F_i$, $G_i$, and $\gamma_{ij}$ by incorporating two criteria, namely, unbiasedness and optimality, on the estimates. For practical implementation, here we give its discrete-time version. First, discretize the target dynamics (1) and the measurement (2) of type-I sensors (see, for example, [1])

\[ x(k) = \Phi x(k - 1) + B w_d(k - 1) \]

\[ y_i(k) = C_j x(k) + v_{i,d}(k), \quad i \in \mathcal{I} \]

where $\Phi \triangleq I_n + \epsilon A$, $\epsilon$ is the time step, and $w_d$ and $v_{i,d}$ are the zero-mean and white noise satisfying

\[ \mathbb{E} \{w_d(k) w_d^T(l)\} = \epsilon^2 Q(d) \]

\[ \mathbb{E} \{v_{i,d}(k) v_{j,d}(l)\} = \epsilon R(d) \]

where $\delta_{kl} = 1$, if $k = l$; and $\delta_{kl} = 0$, otherwise.

Algorithm 1 Algorithm DOCF implemented on sensor $i$

1: Initialization: $\hat{x}_i(0) = \mathbb{E} \{x(0)\}$, $\hat{P}_{ij}(0) = \Pi_i$.
2: loop [Local iteration]
3: \textbf{if} $i \in \mathcal{I}$ \textbf{then}
4: Compute the optimal gain $C_i(k) = \hat{P}_{ij}(k)C_i^T (R_i(k) + C_i \hat{P}_{ij}(k)C_i^T)^{-1}$.
5: Take measurement $y_i(k)$ and update its local estimate $\hat{x}_i(k)$.
6: Compute the perturbed error covariance

\[ \tilde{P}_{ij}(k) = (I_n - G_i(k)C_i) \hat{P}_{ij}(k) (I_n - G_i(k)C_i)^T + \epsilon H_r(k) \sum_{r \in \mathcal{E}_j} a_{ir} (\hat{x}_r(k) - \hat{x}_i(k)) + \epsilon H_r(k) \sum_{r \in \mathcal{E}_j} a_{ir} (\hat{x}_r(k) - \hat{x}_i(k)) H_r^T(k) + G_i(k) R_i(k) G_i^T(k). \]

7: \textbf{else if} $i \in \mathcal{I'}$ \textbf{then}
8: Compute the optimal weights $\tilde{\gamma}_{ij}, j \in \mathcal{I}$ according to (18).
9: Fuse the data received from its neighbors $\tilde{x}_j(k)$
10: Compute the perturbed error covariance $\tilde{P}_{ij}(k)$.
11: \textbf{end if}
12: Update the state of the consensus filter

\[ \hat{x}_i(k) = \Phi \hat{x}_i(k - 1) \]

\[ \tilde{P}_{ij}(k) = \Phi \tilde{P}_{ij}(k - 1) \Phi^T + B Q_k(k - 1) B^T + \epsilon W_i. \]

We assume that all sensors in the network are synchronized so that their communication and estimate updates can be concurrently performed. This way, the discrete-time DOCF is summarized in Algorithm 1. In this algorithm, type-I and type-II sensors can observe the target with dynamics (1) and (2).
IEEE TRANSACTIONS ON CYBERNETICS

IV. MEAN-SQUARE ANALYSIS

In the previous sections, we propose a distributed consensus filter and investigate its properties of unbiasedness and optimality. However, these properties provide no clue about its stability and convergence, which is important for the theoretical analysis and practical application. Here, we give the mean-square analysis of the filter in the Itô stochastic framework.

We first rewrite the target dynamics (1) in the form of an Itô stochastic differential equation [21], i.e.,

\[ dx(t) = Ax(t)dt + B \, d\tilde{w}(t) \quad (19) \]

where \( \tilde{w}(t) \) is an \( m_1 \)-dimensional Brownian motion with \( \mathbb{E}\{d\tilde{w}(t)d\tilde{w}^T(t)\} = Q(t)dt \). In order to obtain a tractable mathematical interpretation of the measurement (2), we introduce a stochastic process \( z_i(t) = \int_0^t y_i(s)ds \) \( \forall i \in \mathcal{I} \). Then, its 407 stochastic representation is given by

\[ dz_i(t) = C_i x(t)dt + d\tilde{v}_i(t) \quad \forall i \in \mathcal{I} \quad (20) \]

where \( \tilde{v}_i(t) \) is an \( m_2 \)-dimensional Brownian motion with \( \mathbb{E}\{d\tilde{v}_i(t)d\tilde{v}_i^T(t)\} = R_{ij}(t)dt \), which is independent of \( x_0 \) and \( \tilde{w}(t) \).

Accordingly, the optimal consensus filter (3) can be rewritten as

\[ d\hat{x}_i(t) = A\hat{x}_i(t)dt + G_i^T(t)dz_i(t) - C_i\hat{x}_i(t)dt \]

\[ + H_i(t) \sum_{j \in \mathcal{N}_i} a_{ij} [\hat{x}_j(t) - \hat{x}_i(t)] dt \quad \forall i \in \mathcal{I}. \quad (21) \]

Denote \( F_i^* \triangleq A - G_i^T C_i \), then subtracting (19) from (21) and using (20) lead to the stochastic representation

\[ dc_i(t) = F_i^*(t)c_i(t)dt + H_i(t) \sum_{j \in \mathcal{N}_i} a_{ij} [e_j(t) - e_i(t)] dt \]

\[ + [-B \, G_i(t)] \begin{bmatrix} d\tilde{w}(t) \\ d\tilde{v}_i(t) \end{bmatrix}. \quad (22) \]

Stack all estimation errors and noise into vectors \( e \triangleq \begin{bmatrix} e_1, e_2, \ldots, e_M \end{bmatrix}^T \) and \( \tilde{v} \triangleq \begin{bmatrix} \tilde{w}, \tilde{v}_1, \ldots, \tilde{v}_M \end{bmatrix}^T \), respectively, and define matrices \( \tilde{P}_i \triangleq [P_{ij}] \in \mathbb{R}^{M \times M} \) and \( R \triangleq [R_{ij}] \in \mathbb{R}^{M \times M} \), then we can obtain the compact vector form of (22) and (14) as follows.

**Lemma 4:** Under the requirement of unbiasedness, the compact forms of the error dynamics (22) and the perturbed covariance dynamics (14) can be rewritten as

\[ dc(t) = \Psi(t)e(t)dt + \Gamma(t)d\tilde{v}(t) \quad (23) \]

\[ \tilde{P}(t) = \Psi(t)\tilde{P}(t) + \Psi(t)\tilde{V}(t) + G^*(t)R(t)G^{*T}(t) \]

\[ + (1_M 1_M^T) \otimes (BQ(t)B^T) + W \quad (24) \]

where \( \Psi = F^* - H(L \otimes I_n) \), \( F^* \triangleq \text{diag}\{F_1, F_2, \ldots, F_M\} \), \( H \triangleq \text{diag}\{H_1, H_2, \ldots, H_M\}, \ G^* \triangleq \text{diag}\{G_1^*, G_2^*, \ldots, G_M^*\} \), \( \Gamma \triangleq [-1_M \otimes B \, G^*] \), \( W \triangleq \text{diag}\{W_1, W_2, \ldots, W_M\} \), and \( L \triangleq [l_{ij}] \in \mathbb{R}^{M \times M} \) with

\[ l_{ij} = \left\{ \begin{array}{ll}
\sum_{j \in \mathcal{N}_i} a_{ij} + \sum_{j \in \mathcal{N}_i \cap \mathcal{I}^c} \sum_{k \in \mathcal{N}_j \cap \mathcal{I}^c} a_{kj} \gamma_{ji}, & j = i \\
-a_{ij} - \sum_{k \in \mathcal{N}_i \cap \mathcal{I}^c} a_{ik} \gamma_{kj}, & j \neq i.
\end{array} \right. \]

**Proof:** See Appendix E.
442 (23) with (24). From (25), we have
\[ \frac{1}{p} \| e \|^2 \leq V(e, t) \leq \frac{1}{2} \| e \|^2. \] (28)

443 Then, utilizing the Itô’s formula [24], we derive that
\[ dV(e, t) = \mathcal{L}V(e, t)dt + 2e^T \tilde{P}^{-1}(t)\Gamma(t)d\tilde{v}(t) \] (29)

where
\[ \mathcal{L}V(e, t) = -e^T \tilde{P}^{-1}(t) \dot{\tilde{P}}(t) \tilde{P}^{-1}(t) + 2e^T \tilde{P}^{-1}(t)\Psi(t)e \]
\[ + \text{tr} \left[ \Gamma(t) \text{diag} \{Q(t), R(t)\} \Gamma(t) \tilde{P}^{-1}(t) \right]. \] (30)

445 With the assumptions (25)–(27) and the expression of \( G_t' \), one obtains
\[ \text{tr} \left[ \Gamma(t) \text{diag} \{Q(t), R(t)\} \Gamma(t) \tilde{P}^{-1}(t) \right] = \text{tr} \left[ (I M T^2) \otimes (BQ(t)B^T) \right. \]
\[ + G^*(t) R(t) G^{**}(t) \tilde{P}^{-1}(t) \leq \frac{1}{p} \left( n M \tilde{\nu} \lambda_{\max}(BB^T) + \tilde{\nu} \sum_{i=1}^M \text{tr} \left[ \tilde{P}_i(t) C_i^T R_i^{-2} C_i \tilde{P}_i(t) \right] \right) \]
\[ \leq \frac{n \bar{\nu} M}{p} \] (31)

447 where \( \lambda_{\max}(\cdot) \) means the largest eigenvalue, and \( \bar{\nu} = \frac{M q \lambda_{\max}(B B^T) + \tilde{\nu}^2}{p} \sum_{i=1}^M \lambda_{\max}(C_i^T C_i) \). As a result, substituting (24) and (31) into (30) leads to
\[ \mathcal{L}V(e, t) \leq -e^T \tilde{P}^{-1}(t)W \tilde{P}^{-1}(t)e + \frac{n \bar{\nu} M}{p} \] (32)

450 where, in the last line, use was made of the positive semidefiniteness of \( BQ(t)B^T \) and \( G^*(t) R(t) G^{**}(t) \).

452 Combining relation (28) and the fact that \( W > 0 \) is positive definite finally enables (32) to be
\[ \mathcal{L}V(e, t) \leq -\frac{\lambda_{\min}(W)}{p^2} \| e \|^2 + \frac{n \bar{\nu} M}{p} \]
\[ \leq -\kappa_1 V(e, t) + \kappa_2 \] (33)

454 where \( \kappa_1 = \frac{n \lambda_{\min}(W)}{p^2} / \tilde{\nu}^2 \), and \( \kappa_2 = n \bar{\nu} M / p \).

455 To complete the proof, we use the stopping time technique and stochastic differential theory [24]. For any given time \( T \geq 0 \) and each positive integer \( k \geq \mathbb{E}[\|e(0)\|] \), define
\[ \tau_{k,T} = \inf \{ t \geq 0 : \| e(t) \| \geq k \} \quad \text{if } \exists t \in [0, T], \| e(t) \| \geq k \]
otherwise.

458 Let \( \tau_{k,T}' \triangleq \min\{t, \tau_{k,T}\} \), then it is apparent that

459 \[ \lim_{k \to \infty} \tau_{k,T}' = t \] almost surely, for all \( 0 \leq t \leq T \).

460 By Itô’s formula and using (33), we have
\[ d \{ e(\exp(\kappa_t) V(e, t)) \} \leq \kappa_2 \exp(\kappa_t) dt \]
\[ + 2 \exp(\kappa_t) e^T \tilde{P}^{-1}(t)\Gamma(t)d\tilde{v}(t). \] (34)

Integrating and then taking expectation of both sides of the preceding equation, we arrive at the following relations:
\[ \mathbb{E} \left\{ V(e(\tau_{k,T}', \tau_{k,T})) \right\} \leq \exp\left(-\kappa_1 \tau_{k,T}'\right) \mathbb{E} \{ V(e(0)) \} \]
\[ + \mathbb{E} \left\{ \int_0^{\tau_{k,T}'} \kappa_2 \exp(\kappa_t(1 - \exp(\kappa_t))) ds \right\} \]
\[ + 2\mathbb{E} \left\{ \int_0^{\tau_{k,T}'} \exp(\kappa_t(1 - \exp(\kappa_t))) e^T(s) \tilde{P}^{-1}(s)\Gamma(s)d\tilde{v}(s) \right\} \]
\[ \leq \mathbb{E} \{ V(e(0)) \} + \frac{\kappa_2}{\kappa_1} \]

where, in the second inequality, use was made of the properties of Itô integral [24]. Note that \( \lim_{k \to \infty} \tau_{k,T}' = t \) almost surely, 464 we can now apply Fatou’s lemma to the preceding inequality to obtain
\[ \mathbb{E} \{ V(e(t)) \} \leq \lim_{k \to \infty} \mathbb{E} \{ V(e(\tau_{k,T}', \tau_{k,T})) \} \]
\[ \leq \mathbb{E} \{ V(e(0)) \} + \frac{\kappa_2}{\kappa_1} \quad \forall t \in [0, T]. \]

It thus follows from (25)–(28), and (31) that there exists a constant \( \mu > 0 \) such that
\[ \mathbb{E} \left\{ \int_0^t \| e(s) \tilde{P}^{-1}(s)\Gamma(s)d\tilde{v}(s) \right\} \leq \mu \mathbb{E} \{ V(e(0)) \} + \frac{\kappa_2}{\kappa_1} \quad \forall t \in [0, T]. \]

Since \( T \) is arbitrary, the properties of Itô integral [24] yield
\[ \mathbb{E} \left\{ \int_0^t e^T(s) \tilde{P}^{-1}(s)\Gamma(s)d\tilde{v}(s) \right\} = 0 \quad \forall t \geq 0. \]

Combining this with (34), we find that
\[ \mathbb{E} \{ V(e(t)) \} \leq \exp(-\kappa_1 t) \mathbb{E} \{ V(e(0)) \} + \frac{\kappa_2}{\kappa_1} (1 - \exp(\kappa_1 t)) \]

which, together with (28), reveals
\[ \mathbb{E} \{ \| e(t) \|^2 \} \leq \frac{\tilde{\nu}}{p} \mathbb{E} \{ \| e(0) \|^2 \} \exp(-\kappa_1 t) + \frac{\kappa_2}{\kappa_1} \tilde{\nu}. \]

Therefore, by Definition 1, we can conclude that the estimation error \( e(t) \) is exponentially bounded in mean square.
detectability properties of the linear system (19). For more details, please refer to [25].

Remark 4: If we only want to obtain the error bounds, $Q(t)$ and $R(t)$ need not to be the covariances of the noise terms. Any positive definite matrices could be applied.

V. SIMULATION STUDY

Here, simulations are presented to verify the efficiency of the proposed DOCF algorithm to track a maneuvering target governed by a Singer model in 2-D space. A Singer model is widely used for modeling maneuvering a target in the 1-D space in the literature of target tracking [1], [26], [27], which assumes that the target acceleration $a(t)$ is modeled as the first-order stationary Markov process

$$\dot{a}(t) = -\alpha a(t) + \omega(t), \quad \alpha > 0 \quad (35)$$

where $\omega(t)$ is the zero-mean white noise with $E\{\omega(t)\omega(\tau)\} = 2\alpha a_0^2 \delta(t - \tau)$, and $\sigma^2_{m_i}$ is the instantaneous variance of $a(t)$.

The state representation of the continuous-time Singer model in the 2-D space can be expressed as

$$\dot{x}(t) = \begin{bmatrix} 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & -\alpha_1 & 0 \\ 0 & 0 & 0 & 0 & 0 & -\alpha_2 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix} x(t) + \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{bmatrix} \omega(t)$$

where $x = [x_1, x_2, x_3, x_4, x_5, x_6]^T$, of which $x_1, x_2$ is the position vector; $[x_3, x_4]^T$ is the velocity vector; $[x_5, x_6]^T$ is the acceleration vector; and $\omega_1(t)$ and $\omega_2(t)$ are the independent jerk noise along the $X$-axis and $Y$-axis, respectively. The initial conditions are $x_0 = [-30, 0, 1, 0.5, 0.2, 0.1]^T$ and $P_0 = I_6$.

Typical values of the parameter $1/\alpha$ for an evasive maneuver are 10–20 s, as suggested in [26]. In the simulations, we choose $\alpha_1 = 0.1$ and $\alpha_2 = 0.05$. In addition, the instantaneous variances $\sigma^2_{m_1}$ and $\sigma^2_{m_2}$ are set to be

$$\sigma^2_{m_1} = \frac{\sqrt{0.27}}{3}[1 + 4 \times 0.2 - 0.3] = 0.135 \text{m}^2/\text{s}^4$$

$$\sigma^2_{m_2} = \frac{\sqrt{0.54}}{3}[1 + 4 \times 0.2 - 0.3] = 0.27 \text{m}^2/\text{s}^4$$

so that $Q = 0.027 I_2$.

We use a sensor network of $N = 20$ sensors consisting of 12 type-I sensors $I = \{1, 2, \ldots, 12\}$ and 8 type-II sensors $I^c = \{13, 14, \ldots, 20\}$, as shown in Fig. 2, to track the maneuvering target. This network is obtained by distributing the sensors randomly over a squared area of 100-m side length. In addition, any two sensors can communicate if the distance between them is smaller than $r = 30$ m. Moreover, each type-I sensor can observe the distorted position of the target according to the 512 linear model (2) with

$$C_i = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \end{bmatrix}, R_i = 0.25 \sqrt{d_{ii}} I_2 \quad \forall i \in I$$

where $d_{ii}$ is the distance between sensor $i$ and the target. The 513 factor $\sqrt{d_{ii}}$ in $R_i$ means that the farther it is from the target, the less information can be observed by the sensor. The other 515 parameters used in the simulations are $H_i = I_6$, $\eta = 2$, and 516 $a_{ij} = \sqrt{1/(1 + d_{ij}^2)}$. Each result presented here is the average 517 of 20 independent runs, except otherwise stated.

Fig. 3 displays the true and estimated trajectories of the target at type-I sensor 2 in one run, from which we can observe that the 520 estimates are close to the true trajectory. This implies that the 521 proposed DOCF algorithm is able to track the target for sensor 522 2. A common sanity test of the estimates is the consistency 523 testing [1], which is crucial for the optimality evaluation. Fig. 3 also demonstrates that the proposed DOCF algorithm produces 525 consistent estimates, in other words, the true target positions 526 are almost within the $3\sigma$ uncertainty bounds centered at the 527 corresponding estimates. The simulation results read that up to 528 91.4% of $x_j$ fall within the interval $[\hat{x}_{2,j} - 3\sqrt{P_2(j,j)}, \hat{x}_{2,j} + 3\sqrt{P_2(j,j)}], j = 1, 2$.

To qualitatively evaluate the performance of the proposed 531 DOCF algorithm, we introduce the two metrics, namely, the 532...
average mean-square deviation (MSD) over all sensors
\[ \text{MSD} = \frac{1}{N} \sum_{i=1}^{N} \| \hat{x}_i - x \|^2 \]
and the average disagreement of the estimates (DoE) among all sensors
\[ \text{DoE} = \frac{1}{N} \sum_{i=1}^{N} \| \hat{x}_i - \hat{x} \|^2 \]
where \( \text{Ave}(\hat{x}) = (1/N) \sum_{i=1}^{N} \hat{x}_i \).

Note that MSD measures the estimation accuracy of the proposed DOCF algorithm, whereas DoE characterizes the differences of the estimates, which reflects the degree of consensus among all sensors. We plot the average MSD and DoE in Fig. 4. It can be seen that both MSD and DoE are small with respect to the measurement noise, which reveals that the proposed DOCF algorithm cannot only track the target with small errors but also possesses the ability of enabling all the sensors to approximately reach agreement on their estimates.

In order to avoid the undesired impact of transient behavior on the estimation errors, we use the average of the root-mean-square error (RMSE) over all sensors to measure the overall performance of the algorithms, which is defined as
\[ \text{RMSE} = \frac{1}{N} \sum_{i=1}^{N} \sqrt{\frac{1}{K-[K/2]+1} \sum_{k=[K/2]}^{K} \| e_i(k) \|^2} \]
where \( K \) is the total number of time steps of simulations, and \( [K/2] \) is the nearest integer around \( K/2 \).

Fig. 5 depicts RMSE-X, RMSE-Y, and RMSE versus the number of type-I sensors, where RMSE-X and RMSE-Y denote the RMSEs for \( X \)-position \( x_1 \) and \( Y \)-position \( x_2 \), respectively. The result demonstrates that the differences of RMSE-X, RMSE-Y, and RMSE are small as the number of type-I sensors varies from 4 to 20, respectively, which indicates that the target tracking problem can be solved using heterogeneous sensor networks. We can further see that the RMSE is the smallest when the number of type-I sensors is 12. This is reasonable because: 1) the measurement noise \( v_i \) is heavily influenced by the distance from the target \( d_{i,t} \) for each type-I sensor \( i \) and 2) only local estimates are communicated between sensors in the DOCF algorithm. If the number of type-I sensors is small, few information about the target could be obtained by the sensors. While if the number of type-I sensors is large, more noise might be injected to the data about the target. Therefore, in both cases, a lower RMSE is certainly not expected. In addition, this suggests that there might be an optimal number of type-I sensors, which is one of our future works to find this optimal value.

Finally, we compare the performance of the proposed DOCF algorithm with the centralized Kalman filter, KCF [(8, Algorithm 3)], diffKF (5, Algorithm 2), and the nonoptimal algorithms. For a fair comparison, we consider the local measurements based DOCF algorithm denoted by DOCF-M, which is given by
\[ \hat{x}_i(t) = (A - G_{i,M}(t)C_{i,M}) \hat{x}_i(t) + G_{i,M}(t) y_{i,M}(t) + E(t) \sum_{j \in N_i} [\hat{x}_j(t) - \hat{x}_i(t)] \]
where \( y_{i,M} = [y_{i1}^T, \ldots, y_{i|N_i|}^T]^T \) is the stacked vector of its neighbors’ measurements and \( C_{i,M} = [C_{i1}^T, \ldots, C_{i|N_i|}^T]^T \). It is straightforward to formally derive the optimal consensus filter in this case following the same line as in the previous 581 sections; thus, we omit the details. For the nonoptimal 582 consensus filter, we set \( G_i(t) = P_i(t)C_i^T R_i^{-1} \) and \( \gamma_i = 1/|N_i| \) \( \forall j \in N_i \).

Since the centralized Kalman filter, KCF, and diffKF algorithms are inherently proposed for homogeneous sensor networks, here, the simulations are performed over the sensor network consisting of only type-I sensors with the same topology as shown in Fig. 2. As for the DOCF-M algorithm, we consider two cases, namely, \( |I^c| = 0 \) and \( |I^c| = 8 \). In Fig. 6, we plot the comparison results with regard to MSD. Clearly, it shows that DoCF-M in the case of \( |I^c| = 0 \) outperforms the 582 other distributed algorithms. Even if \( |I^c| = 8 \) type-II sensors are present in the network, MSD still remains at a satisfactory level, which is lower than that of the KCF algorithm. Moreover, the result demonstrates that the optimal filter DOCF-M possesses improved estimation accuracy MSD compared with the nonoptimal filter.

Additionally, we define an improvement factor (IF) of the compared one
\[ \text{IF} = \frac{\text{RMSE of the compared one} - \text{RMSE of DOCF-M}}{\text{RMSE of the compared one}}. \]
and optimization problems. Furthermore, we have investigated its convergence property. The theoretical analysis has been validated by simulation results that the estimation errors are exponentially bounded. The simulation results also suggest that the heterogeneous sensor network might be a more appropriate choice for the target tracking problem than the homogeneous one.

Some possible directions remain to be further explored, such as the impact of network topology on the accuracy of tracking, further theoretical analysis in the presence of communication delay, and so on. The distributed filter design with no prior knowledge about the process noise of the target is another future work.

VI. CONCLUSION AND FUTURE WORK

We have addressed the distributed tracking problem of a maneuvering target over heterogeneous sensor networks. A novel DOCF is proposed to take the heterogeneity of sensor ability into account by solving the optimal control problems and optimization problems. Furthermore, we have investigated its convergence property. The theoretical analysis has been validated by simulation results that the estimation errors are exponentially bounded. The simulation results also suggest that the heterogeneous sensor network might be a more appropriate choice for the target tracking problem than the homogeneous one.

Some possible directions remain to be further explored, such as the impact of network topology on the accuracy of tracking, further theoretical analysis in the presence of communication delay, and so on. The distributed filter design with no prior knowledge about the process noise of the target is another future work.

APPENDIX A

PROOF OF LEMMA 1

Necessity: Subtracting (1) from (3) and using (4), one obtains for $i \in I$

$$\hat{e}_i(t) = F_i(t)e_i(t) + (F_i(t) + G_i(t)C_i - A)x(t) + G_i(t)v_i(t) + H_i(t)\sum_{j \in N_i} a_{ij} [e_j(t) - \hat{e}_j(t)] - Bw(t) \quad (36)$$

and for $i \in I^c$

$$e_i(t) = \sum_{j \in N_i} \gamma_{ij}(t)e_j(t) + \left(\sum_{j \in N_i} \gamma_{ij}(t) - 1\right)x(t). \quad (37)$$

The requirement of unbiasedness means that $\mathbb{E}\{e_i(t)\} = 0 \forall i \in V$.

Thus, by taking expectation on both sides of (36) and (37), we have $[F_i(t) + G_i(t)C_i - A]\mathbb{E}\{x(t)\} = 0$ and $(\sum_{j \in N_i} \gamma_{ij}(t) - 1)\mathbb{E}\{x(t)\} = 0$. However, in general, $\mathbb{E}\{x(t)\} \not= 0$; it is necessary that $F_i(t) = A - G_i(t)C_i$ in $\mathbb{E}\{x(t)\} = 0$.

Sufficiency: Imposing the assumptions and then taking expectation on both sides of (36), we can obtain

$$\hat{e}_i(t) = F_i(t)e_i(t) + H_i(t)\sum_{j \in N_i} a_{ij} [\hat{e}_j(t) - \hat{e}_i(t)] \quad (38)$$

where $\hat{e}_i(t) = \mathbb{E}\{e_i(t)\}$, for all $i \in V$. Since $\sum_{j \in N_i} \gamma_{ij}(t) = 1$, it follows from (37) and Proposition 1[16] that

$$\hat{e}_i(t) = \sum_{j \in I} \tilde{\gamma}_{ij}(t)\tilde{e}_j(t) \quad (39)$$

where $\tilde{\gamma}_{ij}$ satisfy $\sum_{ij \in E} \tilde{\gamma}_{ij}(t) = 1$. Let $\bar{e}(t) = [\hat{e}_1(t), \hat{e}_2(t), \ldots, \hat{e}_M(t)]^T$, then, by substituting (39) to (38), one gets the linear system $\bar{e}(t) = \Psi(t)\bar{e}(t)$, where $\Psi(t)$ is the coefficient matrix. It is known that the solution is given by $\bar{e}(t) = \Psi(t,0)\bar{e}(0)$, where $\Psi(t,0)$ is the state transition matrix. By the assumption, we have $\bar{e}(0) = 0$. As a result, $\bar{e}(t) = 0$, for all $t > 0$. It thus follows from (39) that $\hat{e}_i(t) = 0 \forall i \in I^c, t > 0$. Therefore, the 661 estimate $\tilde{\bar{e}}_i(t)$ is unbiased for all $i \in V$ and $t > 0$. 662
APPENDIX B
DERIVATION OF (10)–(13)

A) Derivation of (10): Under the unbiased requirement, the estimation error $e_i(t)$ satisfies

$$
\dot{e}_i(t) = (A - G_i(t)C_i)e_i(t) + H_i(t) \sum_{j \in \mathcal{N}_i} a_{ij} [e_j(t) - e_i(t)] + G_i v_i(t) - B w(t) \quad \forall i \in \mathcal{I}.
$$

(40)

The solution $e_i$ to (40) can be expressed as

$$
e_i(t) = \Phi_i(t, 0)e_i(0)
+ \int_0^t \Phi_i(t, \tau) [\psi_i(\tau) + G_i(\tau)v_i(\tau) - Bw(\tau)] d\tau
$$

where $\Phi_i(t, \tau)$ is the state transition matrix corresponding to $A - G_i(t)C_i - H_i(t) \sum_{j \in \mathcal{N}_i} a_{ij}$, and $\psi_i(\tau) \triangleq H_i(t)H_i(\tau)e_i(\tau)$. Since $x(0)$, $w(t)$, and $v_i(t)$, $\forall i \in \mathcal{I}$, are independent, we have

$$
\mathbb{E} \{e_i(t)v_j^T(t)\} = \int_0^t \Phi_i(t, \tau) \mathbb{E} \{\psi_i(\tau)v_j^T(\tau)\} d\tau
$$

where $S_1$

$$
+ \int_0^t \Phi_i(t, \tau)G_i(\tau)\mathbb{E} \{v_i(\tau)v_j^T(\tau)\} d\tau
$$

where $S_2$

$$
\mathbb{E} \{e_i(t)w^T(t)\} = \int_0^t \Phi_i(t, \tau)\mathbb{E} \{\psi_i(\tau)w^T(\tau)\} d\tau - \int_0^t \Phi_i(t, \tau)B\mathbb{E} \{w(\tau)w^T(\tau)\} d\tau.
$$

where $S_3$ and $S_4$

Observe that, for each $r \in \mathcal{N}_i$, $\psi_r(\tau)$ only depends on

$\{e_i(0), v_i(r'), w(r'), r' \leq \tau < t, i \in \mathcal{I}\}$; we find that $S_1 = 0$ and $S_3 = 0$. As for $S_2$ and $S_4$, it follows from (8) and (9), the properties of the Dirac delta function, and the fact that $\Phi_i(t, t) = I_n$ that $S_2 = (1/2)G_i(t)R_i(t)$ and $S_3 = 1/2BQ(t)$. Hence, we have $\mathbb{E} \{e_i(t)v_j^T(t)\} = (1/2)G_i(t)R_i(t)$

(678) and $\mathbb{E} \{e_i(t)w^T(t)\} = -(1/2)BQ(t)$.

(679) Following a similar line, one can obtain $\mathbb{E} \{v_i(t)e_j^T(t)\} = (1/2)R_i(t)G_j^T(t)$ and $\mathbb{E} \{w(t)e_j^T(t)\} = -(1/2)Q(t)B^T$.

Furthermore, we can derive from (40) that

$$
\dot{\hat{P}}_{ij}(t) = \mathbb{E} \{\dot{e}_i(t)e_j^T(t)\} + \mathbb{E} \{e_i(t)e_j^T(t)\}
$$

$$
= (A - G_i(t)C_i)P_{ij}(t) + P_{ij}(t) (A - G_j(t)C_j)^T + H_i(t)P_{N_i,j}(t) + P_{i,N_j}H_j^T(t)
$$

$$
+ G_i(t)\mathbb{E} \{v_i(t)e_j^T(t)\} + \mathbb{E} \{e_i(t)v_j^T(t)\} G_j^T(t)
$$

$$
- BE \{w(t)e_j^T(t)\} - \mathbb{E} \{e_i(t)w^T(t)\} B^T.
$$

which, together with the preceding analysis, yields (10).

Derivation of (11) and (12): We only prove (11) since (12) can be deduced in a similar manner. The unbiasedness reduces (37) to

$$
e_i(t) = \sum_{j \in \mathcal{I}} \tilde{\gamma}_{ij}(t)e_j(t) \quad \forall i \in \mathcal{I}^c
$$

(41)

which implies that, for any two $i \in \mathcal{I}$ and $j \in \mathcal{I}^c$, one can write $\hat{P}_{ij}$ as follows:

$$
\dot{\hat{P}}_{ij}(t) = \mathbb{E} \{\dot{e}_i(t)e_j^T(t)\} + \sum_{l \in \mathcal{I}} \tilde{\gamma}_{il} \mathbb{E} \{e_l(t)e_j^T(t)\}.
$$

Then, repeating the similar arguments as in the derivation of (688) (10) and bearing in mind the constraint (7), we can obtain (11).

B) Derivation of (13): In view of (41), $P_{ij}$, $\forall i, j \in \mathcal{I}$, can be expressed as

$$
P_{ij}(t) = \mathbb{E} \left\{ \sum_{h \in \mathcal{I}} \tilde{\gamma}_{ih} e_h(t) \sum_{l \in \mathcal{I}} \tilde{\gamma}_{jl} e_l^T(t) \right\}
$$

$$
= \sum_{h \in \mathcal{I}} \sum_{l \in \mathcal{I}} \tilde{\gamma}_{ih} \tilde{\gamma}_{jl} P_{hl}(t).
$$

which is just (13).

APPENDIX C
PROOF OF LEMMA 2

In order to compute the optimal gain matrix $G_i$ corresponding to the cost function $J_i(\hat{P}_i)$, we first note that this minimization problem under the constraint (14) is analogous to the classic optimal control problem, where, now, $\hat{P}_i$ can be considered as the state of a system and $G_i$ as the control input. Therefore, the minimum principle of Pontryagin can be used to minimize here as in [28], where it was employed to derive the centralized 701 Kalman–Bucy filter.

To do this, let $t_f$ be the terminal time and the cost function becomes $J_i(\hat{P}_i) = \text{tr} [\hat{P}_i(t_f)]$, $i \in \mathcal{I}$. Then, the optimal control problem with free final state $\hat{P}_i(t_f)$ is readily followed by introducing the Hamiltonian function $H_i(\hat{P}_i, G_i, \Sigma_i, t) = \text{tr} [\hat{P}_i(t)\Sigma_i^T(t)]$, where $\Sigma_i$ is an $n \times n$ matrix of Lagrange multipliers.

According to the minimum principle of Pontryagin [29], the optimal matrix $G_i(t)$ and the corresponding matrix $\Sigma_i(t)$ must 710
satisfy the following conditions:

\[ -\Sigma'_i(t) = \frac{\partial H_i}{\partial P_i} \left( \hat{P}_i^*(t), G_i^*(t), \Sigma_i^*(t), t \right) \]

\[ 0 = \frac{\partial H_i}{\partial G_i^*} \left( \hat{P}_i^*(t), G_i^*(t), \Sigma_i^*(t), t \right) \]

\[ \Sigma_i^*(t_f) = \frac{\partial J_i}{\partial P_i} \left( \hat{P}_i^*(t_f) \right). \]

In view of matrix calculus, it can be derived that

\[ \frac{\partial H_i}{\partial P_i} = \left( A - G_i^*(t)C_i - \sum_{r \in N_i} a_{ir} H_i(t) \right)^T \Sigma_i^*(t) \]

\[ + \Sigma_i^*(t) \left( -\Sigma_i^*(t) \left( A - G_i^*(t)C_i - \sum_{r \in N_i} a_{ir} H_i(t) \right) \right) \]

\[ \frac{\partial H_i}{\partial G_i^*} = -\Sigma_i^*(t) \hat{P}_i^*(t)C_i^T - \sum_{i \in N_i} \Sigma_i^*(t) \hat{P}_i^*(t)C_i^T \]

\[ + \Sigma_i^*(t)G_i^*(t)R_i(t) + \Sigma_i^*(t)G_i^*(t)R_i(t) \]

\[ \frac{\partial J_i}{\partial P_i} = I_n. \]

From the preceding equations, one can obtain

\[ \hat{\Sigma}_i^*(t) = -\left( A - G_i^*(t)C_i - \sum_{r \in N_i} a_{ir} H_i(t) \right)^T \Sigma_i^*(t) \]

\[ -\Sigma_i^*(t) \left( A - G_i^*(t)C_i - \sum_{r \in N_i} a_{ir} H_i(t) \right) \]

with the terminal condition \( \Sigma_i^*(t_f) = I_n \). It then follows from Proposition 1.1[30] that the matrix \( \Sigma_i^*(t) > 0 \), \( \forall t \geq 0 \), and 716 thus nonsingular. In consequence, (43) together with (46) gives

\[ G_i^*(t)R_i(t) = \hat{P}_i^*(t)C_i^T \forall t \geq 0. \]

Since \( R_i(t) \) is assumed to be positive definite and nonnegative, it implies that the optimal gain matrix \( G_i^* \) is given by \( G_i^*(t) = \hat{P}_i^*(t)C_i^TR_i^{-1}(t) \).

The preceding minimization problem can be solved by the method of Lagrangian multipliers. Writing the Kuhn–Tucker conditions, we have

\[ \sum_{i \in N_i} \gamma_i \hat{P}_{ji} \hat{P}_{ji} = 0, \quad k = 2, 3, \ldots, p_i. \]

Note that \( \bar{\gamma}_{ij} \) satisfies the condition (7). From (41), we obtain

\[ \sum_{i \in I} a_{ij} (e_j - e_i) \]

\[ = \sum_{i \in I} a_{ij} (e_j - e_i) + \sum_{i \in I} a_{ij} \gamma_{jk} (e_k - e_i) \]

\[ = \sum_{i \in I} a_{ij} (e_j - e_i) + \sum_{i \in I} \sum_{k \in I} a_{ik} \gamma_{jk} (e_j - e_i) \]

\[ = - (l_i^T \otimes I_n) e \]

Substituting \( \bar{\gamma}_{ij} = 1 - \sum_{h=2}^{p_i} \bar{\gamma}_{ijh} \) into the objective function 725 results in

\[ \min_{\bar{\gamma}_{ijh}} \sum_{h=2}^{p_i} \bar{\gamma}_{ijh} \bar{\gamma}_{ij} \left( \bar{\gamma}_{ijh} \right) \]

\[ + \left( 1 - \sum_{h=2}^{p_i} \bar{\gamma}_{ijh} \right)^2 \bar{\gamma}_{ij} \bar{\gamma}_{ijh} \]

s.t. \( \bar{\gamma}_{ijh} > 0, \quad k = 2, 3, \ldots, p_i. \)

Substituting \( \bar{\gamma}_{ij} = 1 - \sum_{h=2}^{p_i} \bar{\gamma}_{ijh} \) into the objective function 725 results in

\[ \min_{\bar{\gamma}_{ijh}} \sum_{h=2}^{p_i} \bar{\gamma}_{ijh} \bar{\gamma}_{ij} \left( \bar{\gamma}_{ijh} \right) \]

\[ + \left( 1 - \sum_{h=2}^{p_i} \bar{\gamma}_{ijh} \right)^2 \bar{\gamma}_{ij} \bar{\gamma}_{ijh} \]

s.t. \( \bar{\gamma}_{ijh} > 0, \quad k = 2, 3, \ldots, p_i. \)

The preceding minimization problem can be solved by the method of Lagrangian multipliers. Writing the Kuhn–Tucker 728 conditions, we have

\[ \sum_{i \in N_i} \gamma_i \left( \hat{P}_{ji} \hat{P}_{ji} - \hat{P}_{ji} \hat{P}_{ji} \right) + \left( 1 - \sum_{i \in N_i} \gamma_i \right) \left( \hat{P}_{ji} \hat{P}_{ji} - \hat{P}_{ji} \hat{P}_{ji} \right) = 0 \]

and \( \gamma_i > 0 \), for all \( k = 2, 3, \ldots, p_i \). Then, the compact form 730 can be expressed as

\[ U_i \gamma_i = V_i, \quad \text{and} \quad \gamma_i > 0 \]
REFERENCES


Shanying Zhu (S’11) received the B.Sc. degree in information and computing science from North China University of Water Resources and Electric Power, Zhengzhou, China, in 2006 and the M.Sc. degree in applied mathematics from Huazhong University of Science and Technology, Wuhan, China, in 2008. He is currently working toward the Ph.D. degree in the Department of Automation, Shanghai Jiao Tong University, Shanghai, China.

His research interests focus on multiagent systems and wireless sensor networks, particularly in coordination control of mobile robots and distributed detection and estimation in sensor networks and their applications in industrial networks.

Caillian Chen (S’03–M’06) received the B.Eng. and Ph.D. degrees from Carleton University, Canada, in 2003 and 2006, respectively, and the Ph.D. degree in control and systems from City University of Hong Kong, Hong Kong, China, in 2006.

In 2008, she joined the Department of Automation, Shanghai Jiao Tong University, Shanghai, China, as an Associate Professor. She has authored and/or coauthored a research monograph and over 80 referred international journal and conference papers in the areas of intelligent control, nonlinear systems, and wireless sensor 600 networks. She is a holder of six patents. Her current research interests include vehicular ad hoc network, wireless sensor and actuator network, and multiagent and nonlinear systems.

Dr. Chen serves as an Associate Editor of ISRN Sensor Networks and an IPC or TPC member of many conferences, including FUZZ-IEEE, IEEE WCCI, IEEE Globecon, IEEE WISARN, IEEE MSC. She was one of the First Prize Winners of a Natural Science Award from the Ministry of Education of China in 2007. She was a recipient of the "IEEE Transactions on Fuzzy Systems Outstanding Paper Award" in 2008. She was honored as "Shanghai Pujiang Scholar," "Chenguang Scholar," and "SMC Outstanding Young Staff" of Shanghai Jiao Tong University" in 2009.

Wenshuang Li (S’XX) received the B.S. degree in measurement control technology and instruments from Yanshan University, Qinhuangdao, China, in 2010. She is currently working toward the M.S. degree in control science and engineering at Shanghai Jiao Tong University, Shanghai, China.

Her fields of study focus on event-triggered transmission scheme and distributed consensus estimation over wireless sensor networks and their applications in industrial networks.

W. Li is a member of the Center for Intelligent Wireless Networking and Cooperative Control.
Bo Yang (M’XX) received the Ph.D. degree in electrical engineering from the City University of Hong Kong, Kowloon, Hong Kong, in 2009. He was a Postdoctoral Researcher with the Royal Institute of Technology (KTH), Stockholm, Sweden, from 2009 to 2010 and a Visiting Scholar with the Polytechnic Institute of New York University in 2007. In 2010, he joined Shanghai Jiao Tong University, Shanghai, China, where he was first an Assistant Professor and is currently an Associate Professor. He is also a Cyber Scholar with the Cyber Joint Innovation Center founded by Zhejiang University, Hangzhou, China; Tsinghua University, Beijing, China; and Shanghai Jiao Tong University. His current research interests include game theoretical analysis and nonlinear optimization of communication networks and smart grid.

Xinping Guan (M’02–SM’04) received the Ph.D. degree in control and systems from Harbin Institute of Technology, Harbin, China, in 2007. In 2007, he joined the Department of Automation, Shanghai Jiao Tong University, Shanghai, China, where he is currently a Distinguished University Professor and the Director of the Key Laboratory of 911 Systems Control and Information Processing, Ministry of Education of China. He has authored and/or coauthored 2 research monographs, more than 120 SCI indexed papers in IEEE Transactions and other well-known international journals, and numerous conference papers. As a Principal Investigator, he has finished/been working on many national key projects. He is the leader of the prestigious Innovative Research Team of the National Natural Science Foundation of China (NSFC) in 2012. His current research interests include wireless sensor networks, ground–air communication of aircrafts, and cognitive radio networks and their applications in industry.

Dr. Guan is a Committee Member of the Chinese Automation Association Council and the Chinese Artificial Intelligence Association Council. He serves as an Associate Editor of the IEEE TRANSACTIONS ON SYSTEMS, MAN AND CYBERNETICS—PART C; an Editorial Board Committee Member for several Chinese journals; and an International Technical Committee Member in a lot of conferences. He was one of the First Prize Winners of a University Natural Science Award from the Ministry of Education of China in 2003 and 2007, respectively. He was a recipient of the “IEEE Transactions on Fuzzy Systems Outstanding Paper Award” in 2008, the Second Prize of the National Natural Science Award of China, and a “National Outstanding Youth Award” from NSFC. He was appointed as “Changjiang Scholar” by the Ministry of Education of China and “State-level Scholar” of “New Century Bai Qianwan Talent Program” of China.
AUTHOR QUERIES

AUTHOR PLEASE ANSWER ALL QUERIES

A mandatory over length page charge of $175 is required for each page in excess of 12 pages for a regular paper and 6 pages for technical correspondences. There is now a MAXIMUM cap of 2 over length pages.

AQ1 = Information with regard to the prior presentation of the paper was changed to: A preliminary version of this paper was presented at the 8th Asian Control Conference at The Splendor Kaohsiung, Kaohsiung, Taiwan, May 2011. Please check if appropriate. Otherwise, please make the necessary changes.

AQ2 = KCF was considered as the abbreviation for “Kalman-consensus filter” and was thus introduced here. Please check if appropriate. Otherwise, please provide the corresponding expanded form.

AQ3 = The portion “Combining the fact that $W > 0$ is positive definite and relation (28) finally enables (32) to be” was changed to “Combining relation (28) and the fact that $W > 0$ is positive definite finally enables (32) to be...” Please check if appropriate. Otherwise, please make the necessary changes.

AQ4 = Please provide the expanded form of ISRN.

AQ5 = Please provide the expanded forms of IPC and TPC.

AQ6 = Please provide the expanded forms of WCCI and MSC.

AQ7 = Please provide IEEE membership history.

AQ8 = Please provide IEEE membership history.

AQ9 = ACM was expanded as “Association for Computing Machinery.” Please check if appropriate. Otherwise, please provide the corresponding expanded form.

AQ10 = Please provide the expanded form of SCI.

END OF ALL QUERIES
Distributed Optimal Consensus Filter for Target Tracking in Heterogeneous Sensor Networks

Shanying Zhu, Student Member, IEEE, Cailian Chen, Member, IEEE, Wenshuang Li, Student Member, IEEE, Bo Yang, Member, IEEE, and Xinping Guan, Senior Member, IEEE

Abstract—This paper is concerned with the problem of filter design for target tracking over sensor networks. Different from most existing works on sensor networks, we consider the heterogeneous sensor networks with two types of sensors different on processing abilities (denoted as type-I and type-II sensors, respectively). However, questions of how to deal with the heterogeneity of sensors and how to design a filter for target tracking over such kind of networks remain largely unexplored. We propose in this paper a novel distributed consensus filter to solve the target tracking problem. Two criteria, namely, unbiasedness and optimality, are imposed for the filter design. The so-called sequential design scheme is then presented to tackle the heterogeneity of sensors. The minimum principle of Pontryagin is adopted for type-I sensors to optimize the estimation errors. As for type-II sensors, the Lagrange multiplier method coupled with the generalized inverse of matrices is then used for filter optimization. Furthermore, it is proven that convergence property is guaranteed for the proposed consensus filter in the presence of process and measurement noise. Simulation results have validated the performance of the proposed filter. It is also demonstrated that the heterogeneous sensor networks with the proposed filter outperform the homogenous counterparts in light of reduction in the network cost, with slight degradation of estimation performance.

Index Terms—Heterogeneous sensor network, optimal consensus filter, target tracking, unbiased estimate.

I. INTRODUCTION

To locate and track a moving target is crucial for many applications such as robotics, surveillance, monitoring, and security for large-scale complex environments [1]–[3]. In such scenarios, a number of sensors can be employed in order to improve the tracking accuracy and increase the size of the surveillance area in a cooperative manner. Basically, these sensors have modest capabilities of sensing, computation, and multihop wireless communication. Equipped with these 38 capabilities, the sensors can self-organize to form a network that is capable of sensing and processing spatial and temporal dense data in the monitored area.

However, due to radio wave propagation loss in wireless channels, reliable communications between sensors can only be ensured within a short distance. As a result, one of the main challenges for target tracking over sensor networks is that only 43 local information from neighboring sensors is available at each 46 sensor. In this case, the traditional centralized methods (e.g., 47 Kalman filter and extended Kalman filter [1]) for target tracking 48 are not applicable. One possible way is to decentralize the 49 task over the whole network, which brings forward distributed 50 algorithms. The major advantages of distributed algorithms are that: 1) they can reduce the cost of transmitting all data 52 to the fusion center, and 2) they exhibit resilience against 53 sensor failures, thus achieving robustness of the network. A 54 decentralized version of a Kalman filter was proposed in [4], 55 whose main idea is to decentralize the Kalman filtering task to 56 each sensor. However, this algorithm is not scalable with the 57 network size since each sensor needs to communicate with all 58 others to compute its local estimate. A more general setting 59 of the distributed Kalman filter was presented in [5], where 60 a proximity graph is considered, i.e., each sensor only needs to 61 share its information with neighboring sensors. Additionally, a 62 diffusion step is added after the estimate update step to improve 63 the performance of the distributed Kalman filter.

Recently, consensus algorithms of multiagent systems have been proven to be effective tools for performing network-wide distributed tasks [6], [7]. Motivated by these algorithms, Olfati-Saber introduced a scalable distributed Kalman-consensus filter [8] (KCF). It is modified from the Kalman filter by inserting a 69 consensus term to reduce the disagreements of estimates among 70 the sensors. However, it is far from optimal with respect to the 71 error covariance, which might cause unacceptable estimation 72 errors. The optimality and stability analysis in its discrete-time form was further studied in [9]. An alternative distributed 74 Kalman filter based on consensus algorithms was proposed in [10], which is composed of two stages, i.e., a Kalman-like 75 measurement update and an estimate fusion using a consensus 77 matrix. However, only a scalar system was considered, and the 78 optimization of a Kalman gain and a consensus matrix is, in general, nonconvex. In [11], a pinning observer was designed 80 to solve the filtering problem in the case that sensors can only observe partial target information. The distributed filtering 82 with partial information between sensors was investigated 83 in [12].
An optimization approach was also used to solve distributed estimation problems. For example, [13] reformulates the centralized Kalman filter and smoother for distributed operation through the alternating direction method of multipliers. The proposed algorithm offers any-time optimal state estimates based only on local information.

It is noted that most of the aforementioned works deal with homogeneous sensor networks, i.e., all sensors possess identical communication and computation capability, and so on. It is shown that the homogeneous architecture suffers from poor fundamental limits and performance [14]. On the other hand, heterogeneous networks have been experimentally shown to be superior to homogeneous ones due to their potential to increase the network lifetime and reliability [15].

Taking these into considerations, we deal with the target tracking problem in a heterogeneous sensor network framework. Two types of sensors different on processing abilities, which are denoted as type-I and type-II sensors, are present in the network. Type-I sensors are of more computational ability than type-II ones. Questions of how to deal with the heterogeneity of sensors to facilitate the filter design over such kind of sensor networks remain largely unexplored. To the best of the authors’ knowledge, there is no filter design method reported in the open literature for target tracking in such kind of heterogeneous sensor networks. Our main objective is to propose a distributed optimal consensus filter (DOCF) for target tracking relying only on neighbor-to-neighbor communication among sensors. However, the heterogeneity of sensors takes new challenges for the optimal filter design: 1) The coexistence of two types of sensors in the network needs filter forms for different types of sensors. This makes it impossible to design the optimal filter in a unified framework as in [9] and [10], and 13 with identical expressions; 2) The convergence is much more difficult to analyze due to the quite different stochastic representations of the target model and the heterogeneous filter models. The analysis is even untractable without deliberate design of the filter. The main contributions of this paper are twofold:

- We propose a distributed consensus filter for target tracking in heterogeneous sensor networks. The filter parameters are designed according to the performance requirements on unbiasedness and optimality. A sufficient and necessary condition to guarantee the unbiasedness is established. Unlike the homogeneous scenario considered in [9] and [13], we develop the so-called sequential design approach to achieve the optimality with sensor heterogeneity. Specifically, for type-I sensors, the optimization problem is first casted into an optimal control problem. Then, the minimum principle of Pontryagin is applied to solve this problem. Afterward, for type-II sensors, the Lagrange multiplier and least square method is used to obtain the optimal weights. Moreover, the discrete-time DOCF is proposed by discretization and approximation.
- The convergence property of the proposed filter is investigated. We analyze the estimation error dynamics via subtly rewriting the system in the Itô stochastic framework. By using the stochastic Lyapunov method and the stopping time technique, we show that the estimation error of the consensus filter is exponentially bounded in mean square. Simulations have validated the theoretical analysis.

The remainder of this paper is organized as follows: Section II presents the network model of heterogeneous sensor networks, the target model, and the distributed consensus filter. In Section III, we determine the structure of the distributed consensus filter from two aspects, namely, unbiasedness and optimality of the estimates. The mean-square performance of the consensus filter is provided in Section IV. In Section V, simulation results are given to illustrate the effectiveness of the proposed consensus filter with some comparison with some other existing filters. Finally, Section VI concludes this paper.

II. PROBLEM FORMULATION

A. Target and Network Models

The system considered for target tracking problem in this paper is composed of $N$ sensors and a moving target in a $157$ monitored field. The purpose of the sensors is to cooperatively trace the behavior of the target based only on neighbor-to-neighbor communication.

Consider a moving target with the following dynamics:

$$x(t) = Ax(t) + Bu(t)$$

where $A \in \mathbb{R}^{n \times n}$ and $B \in \mathbb{R}^{n \times m}$ are constant matrices, $x(t) \in \mathbb{R}^n$ is the state vector, and $u(t) \in \mathbb{R}^m$ is the ambient $163$ noise, which is assumed to be zero-mean and white. The initial state $x_0$ is a Gaussian random variable with known 165 mean $E\{x_0\} = \bar{x}_0$ and positive definite covariance $E\{(x_0 - \bar{x}_0)(x_0 - \bar{x}_0)^T\} = \Sigma_0 > 0$.

Two types of sensors with different computational capabilities are deployed in the monitored field to locate and track the moving target. Type-I sensors have high processing abilities, whereas type-II sensors are lower-end ones. Furthermore, we assume that only type-I sensors are equipped with onboard 169 sensing units. They can sense the environment and observe the target with noisy measurements of the target. Thus

$$y_i(t) = C_i x(t) + v_i(t) \quad \forall i \in \mathcal{I}$$

where $C_i \in \mathbb{R}^{n_2 \times m_2}$ is the observation matrix, and $v_i(t) \in \mathbb{R}^{m_2}$ is the zero-mean and white measurement noise. As for type-II sensors, they cannot measure the state of the target, which means that all the information about the target at each type-II sensor is directly or indirectly obtained from the type-I sensors.

The sensors in the network are coupled by an undirected 180 communication topology $\mathcal{G} = (\mathcal{V}, \mathcal{E})$, where $\mathcal{V} = \mathcal{I} \cup \mathcal{I}^c = \{1, 2, \ldots, N\}$ is the set of sensors; $\mathcal{I}$ and $\mathcal{I}^c$ are the set of type-I and type-II sensors, respectively; and $\mathcal{E} \subset \mathcal{V} \times \mathcal{V}$ is the set of 183 communication links, over which local information about the target is exchanged. Without loss of generality, let the first $M$ sensors belong to $\mathcal{I}$ and the rest $N-M$ sensors belong to $\mathcal{I}^c$. In this paper, we assume that each sensor transmits at constant 187 power $p_T$ and the receiver has ambient noise power $n_R$. Since 188 path loss is inevitably encountered as the radio wave propagates 189 through the environments, the signal power decreases with the 190 distance captured by the path loss exponent. Then, the signal 191
While the abilities of type-II sensors are much limited, they fuse thus, they can perform more complicated tasks, such as filtering thus scalable with the network size. Fig. 1 depicts a schematic thus the consensus filter (3) and (4) is totally distributed and weighted average in (4). Note that, in the preceding process, to their neighbors. After receiving the local estimates, the estimates from their neighbors. With this information, it is shown to be desirable from a practical viewpoint [18], [19]. 256

As mentioned, the Kalman filter and its distributed general-
207izations for sensor networks [3], [4] are not applicable to target tracking in the heterogeneous sensor networks. We have proposed in [16] a distributed consensus filter for heterogeneous sensor networks.

In this paper, the method in [16] is developed to a novel DOCF for the target tracking scenario. The distributed consensus filter applicable to heterogeneous networks is given as follows. For type-I sensor \( i \in I \)

\[
\dot{x}_i(t) = F_i(t)\dot{x}_i(t) + G_i(t)y_i(t) + H_i(t)\sum_{j \in N_i} a_{ij} \left[ \dot{x}_j(t) - \dot{x}_i(t) \right] \tag{3}
\]

and for type-II sensor \( i \in I^c \)

\[
\dot{x}_i(t) = \sum_{j \in N_i} \gamma_{ij}(t)\dot{x}_j(t) \tag{4}
\]

where \( \dot{x}_i(t) \in \mathbb{R}^n \) is the estimate of the target state \( x(t) \), \( F_i \in \mathbb{R}^{nxn} \) and \( G_i \in \mathbb{R}^{nxm_2} \) are the filter matrices, \( H_i \in \mathbb{R}^{nxn} \)

is the consensus gain matrix, \( a_{ij} = \sqrt{pr/nRd_{ij}^3} \) represents the received signal strength at sensor \( i \) transmitted from sensor \( j \), and \( \gamma_{ij} > 0 \) are constant weights.

Equations (3) and (4) constitute the distributed consensus filter. In the evolution, the type-I sensors sense and measure the location of the target. Meanwhile, they receive the local estimates from their neighbors. With this information, it is possible for the type-I sensors to update their own estimates via (3). Then, the type-I sensors transmit the local estimates to their neighbors. After receiving the local estimates, the type-II sensors scale all the received data in order to find the weighted average in (4). Note that, in the preceding process, communication only takes place among neighboring sensors; thus, the consensus filter (3) and (4) is totally distributed and thus scalable with the network size. Fig. 1 depicts a schematic of the distributed consensus filter at the sensor level.

Remark 1: From (3) and (4), it can be clearly seen that type-I sensors have more powerful computational ability, and thus, they can perform more complicated tasks, such as filtering of its measurement and the local estimates from its neighbors. While the abilities of type-II sensors are much limited, they fuse the data to form their own estimates by only simply weighting all the incoming data.

As for the distributed consensus filter (3) and (4), there are three sets of parameters to be determined, namely, \( F_i \), \( G_i \), and \( \gamma_{ij} \), \( i, j \in V \). To tackle the heterogeneity of (3) and (4), we propose the sequential design approach to determine these parameters in the next section in light of the unbiasedness and optimality of the estimates.

III. STRUCTURE OF CONSENSUS FILTER:

UNBIASEDNESS AND OPTIMALITY

Here, we address the problem of determination of filter matrices \( F_i \), \( G_i \), and \( \gamma_{ij} \) to meet the requirement of unbiasedness, i.e.,

\[
\mathbb{E} \left( e_i(t) \right) = 0, \forall i \in V, t \geq 0,
\]

where \( e_i(t) \) is the estimation error

\[
e_i(t) = \hat{x}_i(t) - x_i(t).
\]

Lemma 1: For each \( i \in V \), let the initial estimate \( \hat{x}_i(0) = \hat{x}_0 \), then \( \hat{x}_i(t), \forall t > 0 \), is unbiased if and only if (1) for type-I sensor, \( F_i(t) = A - G_i(t)C_i \) and 2) for type-II sensor, \( \sum_{j \in N_i} \gamma_{ij}(t) = 1 \).

Proof: See Appendix A.

As a result, it follows from Lemma 1 that the unbiased filter is given by

\[
\hat{x}_i(t) = \left( A - G_i(t)C_i \right) \hat{x}_i(0) + G_i(t)y_i(t)
\]

\[
+ H_i(t)\sum_{j \in N_i} a_{ij} \left[ \dot{x}_j(t) - \dot{x}_i(t) \right].
\]
As for the type-II sensor, from Lemma 1 and Proposition 1 [16], the unbiased filter (4) has the following form:
\[
\hat{x}_i(t) = \sum_{j \in \mathcal{I}} \hat{\gamma}_{ij}(t) \hat{x}_j(t), \quad i \in \mathcal{I}^c
\] (6)

where \( \mathcal{I} = \{i \in \mathcal{I} : \text{there exists } j \in \mathcal{I}^c \text{ such that } (i,j) \in \mathcal{E} \} \) and we can rearrange the sensors such that \( j_k : \mathcal{I} \rightarrow \mathcal{I} \), i.e.,
\[
\hat{\gamma}_{ijk}(t) = \begin{cases} 0, & k = 1, \ldots, p_i, \\ \hat{\gamma}_{ijk} + 1, & k = p_i + 1, \ldots, |\mathcal{I}|, \end{cases}
\]
and \( \sum_{k=1}^{p_i} \hat{\gamma}_{ijk}(t) = 1. \) (7)

### 2.2 Optimization of the Gain Matrix \( G_i \) and Weights \( \hat{\gamma}_{ij} \)

Optimal filters possess some important merits, such as being robust in their maintenance of performance standards [18]. In the target tracking scenario, optimal filters for homogeneous sensor networks were considered in [9] and [13]. In our setting, however, the heterogeneity of sensors presents a challenge for the optimal filter design. We propose the sequential design approach to tackle this difficulty.

To facilitate the filter design, we give some standard assumptions. The random variables \( x_0, w(t), \) and \( v_i(t), i \in \mathcal{I}, \) are 282 independent, and
\[
\begin{align*}
E \{w(t)w(t)^T\} &= Q(t) \delta(t - \tau) \quad \forall \tau \in \mathcal{V} \\
E \{v_i(t)v_j^T(t)\} &= R_{ij}(t) \delta(t - \tau) \quad \forall \tau \in \mathcal{V}
\end{align*}
\] (8)

where \( \delta(\cdot) \) is the Dirac delta function. Moreover, assume that
\[
R_i(t) \triangleq R_{ii}(t), \quad \forall i \in \mathcal{I}, \text{ is positive definite.}
\]

Define the error covariance matrix as
\[
P_{ij}(t) \triangleq E\{e_i(t)e_j^T(t)\}, \quad \forall i,j \in \mathcal{V}, \quad \text{and denote}
\]
\[
P_{\mathcal{N}_i \setminus \mathcal{I}^c}(t) \triangleq \sum_{j \in \mathcal{N}_i \setminus \mathcal{I}^c} a_{ij} [P_{ij}(t) - P_{ij}^*(t)] \quad \text{and } P_{\mathcal{N}_i \setminus \mathcal{I}^c}(t) \triangleq \sum_{j \in \mathcal{N}_i \setminus \mathcal{I}^c} a_{ij} P_{ij}(t).
\]

Then, a standard argument shows that \( P_{ij}(t) \) satisfies (see Appendix B) the cases here.

#### Case i) \( i \in \mathcal{I} \) and \( j \in \mathcal{I} \).

\[
\begin{align*}
P_{ij}(t) &= (A - G_i(t)C_i) P_{ij}(t) + P_{ij}(t) (A - G_j(t)C_j)^T \\
&\quad + H_i(t) R_{\mathcal{N}_i \setminus \mathcal{I}^c}(t) H_j^T(t) \\
&\quad + G_i(t) R_{ij}(t) G_j^T(t) + BQ(t)B^T.
\end{align*}
\] (9)

#### Case ii) \( i \in \mathcal{I} \) and \( j \in \mathcal{I}^c \).

\[
\begin{align*}
P_{ij}(t) &= (A - G_i(t)C_i) P_{ij}(t) \\
&\quad + \sum_{k \in \mathcal{I}} \hat{\gamma}_{ik}(t) P_{ik}(t) (A - G_k(t)C_k)^T \\
&\quad + H_i(t) R_{\mathcal{N}_i \setminus \mathcal{I}^c}(t) H_k^T(t) \\
&\quad + \sum_{k \in \mathcal{I}} \hat{\gamma}_{ik}(t) G_i(t) R_{ki}(t) G_k^T(t) + BQ(t)B^T.
\end{align*}
\] (10)

#### Case iii) \( i \in \mathcal{I}^c \) and \( j \in \mathcal{I} \).

\[
\begin{align*}
\hat{P}_{ij}(t) &= P_{ij}(t) (A - G_j(t)C_j)^T \\
&\quad + \sum_{k \in \mathcal{I}} \hat{\gamma}_{ik}(t) (A - G_k(t)C_k) P_{kj}(t) \\
&\quad + P_{\mathcal{N}_i \setminus \mathcal{I}^c}(t) H_j^T(t) + \sum_{k \in \mathcal{I}} \hat{\gamma}_{ik}(t) H_k(t) P_{\mathcal{N}_i \setminus \mathcal{I}^c}(t) \\
&\quad + \sum_{k \in \mathcal{I}} \hat{\gamma}_{ik}(t) G_k(t) R_{kj}(t) G_j^T(t) + BQ(t)B^T. \quad \text{(12)}
\end{align*}
\]

#### Case iv) \( i \in \mathcal{I}^c \) and \( j \in \mathcal{I}^c \).

\[
\begin{align*}
P_{ij}(t) &= \sum_{k \in \mathcal{I}^c} \sum_{l \in \mathcal{I}^c} \hat{\gamma}_{ik} \hat{\gamma}_{jl} P_{kl}(t).
\end{align*}
\] (13)

From a physical point of view, it is desirable to have an unbiased estimate with minimum covariance \( \hat{P}_i \triangleq \hat{P}_i \). In the following, however, we consider the perturbed covariance matrix \( \hat{P}_i \) in view of convergence performance:

\[
\begin{align*}
\hat{P}_i(t) &= (A - G_i(t)C_i) \hat{P}_i(t) + \hat{P}_i(t) (A - G_i(t)C_i)^T \\
&\quad + H_i(t) R_{\mathcal{N}_i \setminus \mathcal{I}^c}(t) H_i^T(t) \\
&\quad + G_i(t) R_{ij}(t) G_i^T(t) + BQ(t)B^T + W_i(t)
\end{align*}
\] (14)

where \( W_i \) is positive definite, \( \hat{P}_{\mathcal{N}_i \setminus \mathcal{I}^c}(t) \triangleq \sum_{j \in \mathcal{N}_i \setminus \mathcal{I}^c} a_{ij} [\hat{P}_{ij}(t) - \hat{P}_{ij}^*(t)] \) and \( \hat{P}_{ij}^*(t) \) for all \( j \in \mathcal{V} \) are likewise determined as \( P_{ij}^*(t) \) in four cases. Denote 300 the corresponding equations as (10'), (11'), (12'), and (13'), 331 respectively. Moreover, we set \( \hat{P}_i(0) = P_{ij}(0) = 0 \).

We introduce the cost function
\[
J_i(\hat{P}_i) = \text{tr} \left[ \hat{P}_i(t) \right], \quad i \in \mathcal{V}
\] (15)
as the measure of the performance of the filter, where \( \text{tr}[\cdot] \) denotes the trace of a matrix. The smaller the value of \( J_i \), the better the consensus filter (5) and (6).

It is observed from (10)–(13) that, with the cost function \( J_i \), matrix \( G_i \) can be independently determined from weights \( \hat{\gamma}_{ij} \) for type-I and type-II sensors, respectively. We first focus on the selection of \( G_i \) for the type-I sensors.

**Lemma 2:** For type-I sensor \( i \in \mathcal{I} \), the optimal matrix \( G_i^* \) with respect to the cost function \( J_i(\hat{P}_i) \) is given by
\[
\hat{P}_i(t) = A \hat{P}_i(t) + \hat{P}_i(t)A^T - \hat{P}_i(t)C_i^T R_i^{-1}(t)C_i \hat{P}_i(t)
\]
\[
+ H_i(t) \hat{P}_{\mathcal{N}_i \setminus \mathcal{I}^c}(t) + \hat{P}_{\mathcal{N}_i \setminus \mathcal{I}^c}(t) H_i^T(t)
\]
\[
+ BQ(t)B^T + W_i(t).
\] (16)
TABLE I

<table>
<thead>
<tr>
<th>Unbiasedness</th>
<th>type-I sensor</th>
<th>type-II sensor</th>
</tr>
</thead>
<tbody>
<tr>
<td>F(t) = A - G_i(t)C_i</td>
<td>( \sum_{j \in I} \gamma_{ij} = 1 )</td>
<td>( \sum_{j \in I} \gamma_{ij} = 1 )</td>
</tr>
<tr>
<td>G_i(t) = ( \hat{P}_i(t)C_i^T \hat{R}^{-1} )</td>
<td>( \hat{z}<em>{ij}^* = 1 - \sum</em>{k=1}^{p_i} z_{jk}, \ j = j_1, j_2, \ldots, j_{p_i} )</td>
<td>( \hat{z}<em>{ij}^* = z_j, \ \forall j = j_2, j_3, \ldots, j</em>{p_i}, \text{and otherwise 0} )</td>
</tr>
</tbody>
</table>

Proof: See Appendix C.

With the optimal \( G_i \), we proceed to the optimization of \( E \) weights \( \gamma_{ij} \) for type-II sensor \( i \) such that \( J_i(\hat{P}_i) \) is minimized.

\[
\gamma_{ij}^* = \begin{cases} 
1 - \sum_{k=1}^{p_i} z_{jk}, & j = j_1, \\
0, & j = j_2, \ldots, j_{p_i},
\end{cases}
\tag{17}
\]

where \( z \equiv [z_{j_2}, z_{j_3}, \ldots, z_{j_{p_i}}]^T > 0 \) is the positive solution of the linear equation \( U_i z = V_i \), in which

\[
U_i = \begin{bmatrix} 
U_{j_2,j_2} & \cdots & U_{j_2,j_{p_i}} \\
\vdots & \ddots & \vdots \\
U_{j_{p_i},j_2} & \cdots & U_{j_{p_i},j_{p_i}} 
\end{bmatrix}, \quad V_i = \begin{bmatrix} 
V_{j_2} \\
\vdots \\
V_{j_{p_i}} 
\end{bmatrix}
\]

and \( U_{j_k,j_l} = \text{tr}[\hat{P}_{j_k,j_l} - \hat{P}_{j_k,j_l}] \) and \( V_{j_k} = \text{tr}[\hat{P}_{j_k,j_k}] \) for \( k, l = 2, 3, \ldots, p_i \).

Proof: See Appendix D.

In general, there may exist no positive solution or more than one positive solution of \( U_i z = V_i \). In this case, we turn to the least squares solution by minimizing the norm of the residual, i.e.,

\[
\min_{z \in \mathbb{R}^p} \| U_i z - V_i \|, \quad \text{s.t.} \quad z > 0.
\tag{18}
\]

Several numerical methods can be used here to solve this problem efficiently such as the active set algorithm [20].

Summarizing Lemmas 1, 2, and 3, we can finally determine the parameters of the consensus filter at each sensor as follows.

**Theorem 1:** Consider a heterogeneous sensor network with communication topology \( G \) tracking the target with dynamics (1) and (3) (4) is optimal with respect to the cost function \( J_i(\hat{P}_i) \) and (15) with the parameters given in Table I.

**C. Discrete-Time Version of the Consensus Filter: DOCF**

In the previous subsections, we have demonstrated how to select the parameters \( F_i, G_i, \) and \( \gamma_{ij} \) by incorporating two criteria, namely, unbiasedness and optimality, on the estimates. For practical implementation, here we give its discrete-time version. First, discretize the target dynamics (1) and the measurement (2) of type-I sensors (see, for example, [1])

\[
x(k) = \Phi x(k - 1) + B w_d(k - 1) \\
y_i(k) = C_i x(k) + v_{i,d}(k), \quad i \in I
\]

where \( \Phi = I_n + \epsilon A, \) \( \epsilon \) is the time step, and \( w_d \) and \( v_{i,d} \) are the zero-mean and white noise satisfying

\[
E \{ w_d(k)w_d^T(l) \} = Q(k) \delta_{kl} \Delta Q_d(k) \delta_{kl} \\
E \{ v_{i,d}(k)v_{j,d}^T(l) \} = R_{ij}(k) \delta_{ij} \Delta R_{ij}(k) \delta_{ij}
\]

where \( \delta_{kl} = 1, \) if \( k = l; \) and \( \delta_{kl} = 0, \) otherwise.

**Algorithm 1** Algorithm DOCF implemented on sensor \( i \)

1: Initialization: \( \hat{x}_i(0) = E[x(0)], \hat{P}_i(0) = \hat{P}_{ii}(0) = I_0. \)
2: loop [Local iteration]
3: \( \text{if } i \in I \text{ then} \)
4: \( \text{Compute the optimal gain } G_i(k) = \hat{P}_i(k)C_i^T (R_i^d(k) + C_i \hat{P}_i(k)C_i^T \delta_{ij} \Delta R_{ij}(k))^{-1}. \)
5: \( \text{Take measurement } y_i(k) \) and update its local estimate \( \hat{x}_i(k) = \hat{x}_i(k) + \Delta \hat{P}_i(k) \)
6: Compute the perturbed error covariance \( \hat{P}_i(k) = (I_n - G_i(k)C_i) \hat{P}_i(k) (I_n - G_i(k)C_i)^T + \Delta \hat{P}_i(k). \)
7: \( \text{else if } i \in \bar{I} \text{ then} \)
8: \( \text{Compute the optimal weights } \gamma_{ij}, j \in \bar{I} \text{ according to (18)} \)
9: \( \text{Fuse the data received from its neighbors } \hat{x}_j(k) = \sum_{j \in \bar{I}} \gamma_{ij} \hat{x}_j(k). \)
10: Compute the perturbed error covariance \( \hat{P}_i(k) = \sum_{j \in \bar{I}} \gamma_{ij} \hat{P}_j + \hat{P}_{ii}(k). \)
11: \( \text{end if} \)
12: Update the state of the consensus filter \( \hat{x}_i(k) = \Phi \hat{x}_i(k - 1) \)
13: \( \hat{P}_i(k) = \Phi \hat{P}_i(k - 1) \Phi^T + BQ^d(k - 1)B^T + \epsilon W_i. \)
14: \( \text{end loop} \)

We assume that all sensors in the network are synchronized so that their communication and estimate updates can be concurrently performed. This way, the discrete-time DOCF is summarized in Algorithm 1. In this algorithm, type-I and type-II sensors cooperate to estimate the target state by solving the optimization problem (17) in parallel.
373 type-II sensors are treated separately due to the sequential 374 design approach. For each type-I sensor, it first computes the 375 optimal gain $G_i$, broadcasts it to all its neighbors, and then up- 376dates its estimates (lines 4 and 5). As for the estimate update, it 377fuses its measurement and the local estimates received from its 378neighbors (line 5). Afterward, it calculates the perturbed error 379covariance $P_{ij}$ based on the received data from its neighbors 380(line 6). On the other hand, for each type-II sensor, it only in- 381corporates all its neighbors’ estimates (line 9) and the perturbed 382error covariance (line 10) based on the optimal weights $\tilde{\gamma}_{ij}$ 383(line 8).

384 Remark 2: It is noted that, in Algorithm 1, we make some 385approximations of $G_i$ and $P_{ij}$ by neglecting the $o(e)$ and $o(e^2)$ 386terms during discretization, respectively (see lines 4 and 6). 387Moreover, it requires that, at every iteration, sensor $i$ should 388transmit the prior estimate $\hat{x}_i$, the corresponding covariance 389matrix $\tilde{P}_{ij}$, the optimal gain $G_i$, and the measurement matrix 390$C_i$ to its neighbors, and even its second-hop neighbors (see line 3916). It might become a heavy burden for the sensors with limited 392energy. In this case, a suboptimal alternative of the update of 393$P_{ij}$ could be

$$\tilde{P}_{ij}(k) = (I_n - G_i(k)C_i) \tilde{P}_{ij}(k) (I_n - G_j(k)C_j)^T + G_i(k)R_{ij}(k)G_j(k).$$

IV. MEAN-SQUARE ANALYSIS

395 In the previous sections, we propose a distributed consensus 396filter and investigate its properties of unbiasedness and opti- 397mality. However, these properties provide no clue about its 398stability and convergence, which is important for the theoretical 399analysis and practical application. Here, we give the mean- 400square analysis of the filter in the Itô stochastic framework.

401 We first rewrite the target dynamics (1) in the form of an Itô 402stochastic differential equation [21], i.e.,

$$dx(t) = Ax(t)dt + Bdw(t)$$  

403where $\bar{w}(t)$ is an $m_1$-dimensional Brownian motion with 404$E\{\bar{w}(t)\bar{w}^T(t)\} = Q(dt)$. In order to obtain a tractable 405mathematical interpretation of the measurement (2), we intro- 406duce a stochastic process $z_i(t) = \int_0^t y_i(s)ds \forall i \in \mathcal{I}$. Then, its 407stochastic representation is given by

$$dz_i(t) = C_i x(t)dt + d\bar{v}_i(t) \quad \forall i \in \mathcal{I}$$  

408where $\bar{v}_i(t)$ is an $m_2$-dimensional Brownian motion with 409$E\{\bar{v}_i(t)\bar{v}_i^T(t)\} = R_{ij}(dt)$, which is independent of $x_0$ 410and $\bar{w}(t)$.

411 Accordingly, the optimal consensus filter (3) can be 412rewritten as

$$d\hat{x}_i(t) = A\hat{x}_i(t)dt + G_i^T(t)dz_i(t) - C_i \hat{x}_i(t)dt$$

$$+ H_i(t) \sum_{j \in \mathcal{N}_i} a_{ij} [\hat{x}_j(t) - \hat{x}_i(t)] dt \quad \forall i \in \mathcal{I}.$$  

414Using (20) lead to the stochastic representation

$$de_i(t) = F_i^t(t)e_i(t)dt + H_i(t) \sum_{j \in \mathcal{N}_i} a_{ij} [e_j(t) - e_i(t)] dt + [-B \quad G_i^T(t)] \begin{bmatrix} d\bar{v}(t) \\ d\bar{v}_i(t) \end{bmatrix}. \quad (22)$$

416 Stack all estimation errors and noise into vectors $e \triangleq 417\begin{bmatrix} e_1, e_2, \ldots, e_M \end{bmatrix}^T$ and $\bar{v} \triangleq 418\begin{bmatrix} \bar{w}, \bar{v}_1, \ldots, \bar{v}_M \end{bmatrix}^T$, respectively, and define matrices $\bar{P} \triangleq \begin{bmatrix} \bar{P}_{ij} \end{bmatrix} \in \mathbb{R}^{M \times M}$ and $R \triangleq \begin{bmatrix} R_{ij} \end{bmatrix} \in \mathbb{R}^{M \times M}$, then we can obtain the compact vector form of (22) and (14) as follows.

419 Lemma 4: Under the requirement of unbiasedness, the comp- 420act forms of the error dynamics (22) and the perturbed covari- 421ance dynamics (14) can be rewritten as

$$de(t) = \Psi(t)e(t)dt + \Gamma(t)d\bar{v}(t) \quad (23)$$

$$\bar{P}(t) = \Psi(t)\bar{P}(t) + \bar{P}(t)\Psi^T(t) + G^*(t)R(t)G^T(t) + (1_M1_M^T) \otimes (BQ(t)B^T) + W \quad (24)$$

425 where $\Psi = F^* - H(L \otimes I_n)$, $F^* \triangleq \text{diag}\{F_1, F_2, \ldots, F_M\}$, $H \triangleq \text{diag}\{H_1, H_2, \ldots, H_M\}$, $G^* \triangleq \text{diag}\{G_1, G_2, \ldots, G_M^*\}$, $\Gamma \triangleq [-1_M \otimes B \quad G^*]$, $W \triangleq \text{diag}\{W_1, W_2, \ldots, W_M\}$, and $L \triangleq [l_{ij}] \in \mathbb{R}^{M \times M}$ with

$$l_{ij} = \begin{cases} \sum_{j \in \mathcal{N}_i \cup \mathcal{I}} a_{ij} + \sum_{j \in \mathcal{N}_i \cup \mathcal{I}} \sum_{k \in \mathcal{I}} a_{ik} \gamma_{jk} & j = i \\ -a_{ij} - \sum_{k \in \mathcal{I}} a_{ik} \gamma_{jk} & j \neq i. \end{cases}$$

426

427 IV. MEAN-SQUARE ANALYSIS

428

429 In order to investigate the mean-square performance of the 428 error dynamics (23) with (24), we adopt the following definition 429of stochastic stability (see, for example, [22] and [23]).

430 Definition 1: The stochastic process $e(t)$ is said to be ex- 431ponentially bounded in mean square if there exist constants $\beta_i > 0$, $i = 1, 2, 3$ such that

$$E\{\|e(t)\|^2\} \leq \beta_i E\{\|e(0)\|^2\} \exp(-\beta_2 t) + \beta_3 \quad \forall t \geq 0.$$  

432

433

434 We are now ready to give the following theorem.

435 Theorem 2: Consider the stochastic differential equations 435 (23) and (24). Suppose that $\bar{P}(t)$, $Q(t)$, and $R(t)$ are bounded, i.e., there exist positive constants $p, \bar{p}, q, \bar{q}, \bar{r}, \bar{r} > 0$ such that

$$l_{1M} \leq \bar{P}(t) \leq \bar{p}l_{1M}$$  

$$Q(t) \leq \bar{q}I_n$$  

$$l_{1M} \leq R(t) \leq \bar{r}l_{1M}$$  

then the estimation error $e(t)$ is exponentially bounded in mean 438 square.

439 Proof: Define the stochastic process $V(e, t) = 440 e^T \bar{P}(t)e$ corresponding to the estimation error dynamics 441
Then, utilizing the Itô’s formula [24], we derive that
\[ dV(e, t) = LV(e, t)dt + 2e^T \tilde{P}^{-1}(t) \Gamma(t) d\tilde{v}(t) \] (29)
where
\[ LV(e, t) = -e^T \tilde{P}^{-1}(t) \dot{\hat{P}}(t) \tilde{P}^{-1}(t) + 2e^T \tilde{P}^{-1}(t) \Psi(t)e \]
\[ + \text{tr} \left[ \Gamma(t) \text{diag} \{ Q(t), R(t) \} \Gamma(t) \tilde{P}^{-1}(t) \right]. \] (30)

With the assumptions (25)–(27) and the expression of \( G_t \), one obtains
\[ \text{tr} \left[ \Gamma(t) \text{diag} \{ Q(t), R(t) \} \Gamma(t) \tilde{P}^{-1}(t) \right] \]
\[ = \text{tr} \left[ \left( M I^T M \right) \otimes \left( B Q(t) B^T \right) \right. \]
\[ + G^*(t) R(t) G^*(t) \tilde{P}^{-1}(t) \]
\[ \leq \frac{1}{P} \left( n M \tilde{q} \lambda_{\text{max}}(B B^T) + \bar{r} + \sum_{i=1}^{M} \text{tr} \left[ \tilde{P}_i(t) C_i^T R_i^{-2} C_i \tilde{P}_i(t) \right] \right) \]
\[ \leq \frac{n \theta_M}{P}. \] (31)

where \( \lambda_{\text{max}}(\cdot) \) means the largest eigenvalue, and \( \theta_M = M \tilde{q} \lambda_{\text{max}}(B B^T) + \bar{r} + \sum_{i=1}^{M} \lambda_{\text{max}}(C_i^T C_i) \). As a result, substituting (24) and (31) into (30) leads to
\[ LV(e, t) \leq -e^T \tilde{P}^{-1}(t) W \tilde{P}^{-1}(t) e + \frac{n \theta_M}{P}. \] (32)

where, in the last line, use was made of the positive semidefiniteness of \( B Q(t) B^T \) and \( G^*(t) R(t) G^*(t) \).

Combining relation (28) and the fact that \( W > 0 \) is positive definite finally enables (32) to be
\[ LV(e, t) \leq -\frac{\lambda_{\text{min}}(W)}{\bar{p}^2} \| e \|^2 + \frac{n \theta_M}{P} \]
\[ \leq -\kappa_1 V(e, t) + \kappa_2 \] (33)

where \( \kappa_1 = \frac{\lambda_{\text{min}}(W)}{\bar{p}^2} \), and \( \kappa_2 = \frac{n \theta_M}{P} \).

To complete the proof, we use the stopping time technique of stochastic differential theory [24]. For any given time \( T \geq 0 \) and each positive integer \( k \geq \mathbb{E}(\| e(0) \|) \), define
\[ \tau_{k, T} = \inf \left\{ t \geq 0 : \| e(t) \| \geq k \right\}, \quad \text{if } \exists t \in [0, T], \| e(t) \| \geq k \]
\[ = T, \] otherwise.

Let \( \tau_{k, T} \hat{=} \min\{t, \tau_{k, T}\} \), then it is apparent that
\[ \lim_{k \to \infty} \tau_{k, T} = t \] almost surely, for all \( 0 \leq t \leq T \).

By Itô’s formula and using (33), we have
\[ d (\exp(\kappa_1 t) V(e, t)) \leq \kappa_2 \exp(\kappa_1 t) dt \]
\[ + 2 \exp(\kappa_1 t) e^T \tilde{P}^{-1}(t) \Gamma(t) d\tilde{v}(t). \] (34)

Integrating and then taking expectation of both sides of the preceding equation, we arrive at the following relations:
\[ \mathbb{E} \left\{ V(e, \tau_{k, T}) \right\} \leq \exp(-\kappa_1 \tau_{k, T}) \mathbb{E} \left\{ V(e(0), 0) \right\} \]
\[ + \mathbb{E} \left\{ \int_0^{\tau_{k, T}} \kappa_2 \exp(\kappa_1(s - \tau_{k, T})) ds \right\} \]
\[ + 2 \mathbb{E} \left\{ \int_0^{\tau_{k, T}} \exp(\kappa_1(s - \tau_{k, T})) e^T(s) \tilde{P}^{-1}(s) \Gamma(s) d\tilde{v}(s) \right\} \]
\[ \leq \mathbb{E} \left\{ V(e(0), 0) \right\} + \frac{\kappa_2}{\kappa_1} \]

where, in the second inequality, use was made of the properties of Itô integral [24]. Note that \( \lim_{k \to \infty} \tau_{k, T} = t \) almost surely, we can now apply Fatou’s lemma to the preceding inequality to obtain
\[ \mathbb{E} \left\{ V(e, t) \right\} \leq \lim_{k \to \infty} \mathbb{E} \left\{ V(e, \tau_{k, T}) \right\} \]
\[ \leq \mathbb{E} \left\{ V(e(0), 0) \right\} + \frac{\kappa_2}{\kappa_1} < \infty \quad \forall t \in [0, T]. \]

It thus follows from (25)–(28), and (31) that there exists a constant \( \mu > 0 \) such that
\[ \mathbb{E} \left\{ \int_0^t e^T(s) \tilde{P}^{-1}(s) \Gamma(s) d\tilde{v}(s) \right\} \]
\[ \leq \mu \bar{p} \mathbb{E} \left\{ \int_0^t V(e(s), s) ds \right\} \]
\[ \leq \mu \bar{p} T \mathbb{E} \left\{ V(e(0), 0) \right\} + \frac{\kappa_2}{\kappa_1} < \infty \quad \forall t \in [0, T]. \]

Since \( T \) is arbitrary, the properties of Itô integral [24] yield
\[ \mathbb{E} \left\{ \int_0^t e^T(s) \tilde{P}^{-1}(s) \Gamma(s) d\tilde{v}(s) \right\} = 0 \quad \forall t \geq 0. \]

Combining this with (34), we find that
\[ \mathbb{E} \left\{ V(e, t) \right\} \leq \exp(-\kappa_1 t) \mathbb{E} \left\{ V(e(0), 0) \right\} + \frac{\kappa_2}{\kappa_1} (1 - \exp(\kappa_1 t)) \]

which, together with (28), reveals
\[ \mathbb{E} \left\{ \| e(t) \|^2 \right\} \leq \frac{\bar{p}}{P} \mathbb{E} \left\{ \| e(0) \|^2 \right\} \exp(-\kappa_1 t) + \frac{\kappa_2 \bar{p}}{\kappa_1}. \]

Therefore, by Definition 1, we can conclude that the estimation error \( e(t) \) is exponentially bounded in mean square.

Remark 3: In Theorem 2, bounds of \( \tilde{P} \) are required to prove the exponential boundedness of the estimation error \( e(t) \). The condition (25) is closely related with the observability and
detectability properties of the linear system (19). For more details, please refer to [25].

**Remark 4:** If we only want to obtain the error bounds, \(Q(t)\) and \(R(t)\) need not to be the covariances of the noise terms. Any positive definite matrices could be applied.

V. SIMULATION STUDY

Here, simulations are presented to verify the efficiency of the proposed DOCF algorithm to track a maneuvering target governed by a Singer model in 2-D space.

A Singer model is widely used for modeling maneuvering a target in the 1-D space in the literature of target tracking [1], [26], [27], which assumes that the target acceleration \(a(t)\) is modeled as the first-order stationary Markov process

\[
\dot{a}(t) = -\alpha a(t) + \omega(t), \quad \alpha > 0 
\]

where \(\omega(t)\) is the zero-mean white noise with \(E\{\omega(t)\omega(\tau)\} = 2\alpha \sigma_m^2 \delta(t - \tau)\), and \(\sigma_m^2\) is the instantaneous variance of \(a(t)\).

The state representation of the continuous-time Singer model in 2-D space can be expressed as

\[
\dot{x}(t) = \begin{bmatrix}
  0 & 0 & 0 & 0 & 0 \\
  0 & 0 & 0 & 0 & 0 \\
  0 & 0 & 0 & 0 & 1 \\
  0 & 0 & 0 & 0 & 0 \\
  0 & 0 & 0 & 0 & 0 \\
\end{bmatrix} x(t) + \begin{bmatrix}
  \omega_1(t) \\
  \omega_2(t) \\
\end{bmatrix}
\]

where \(x = [x_1, x_2, x_3, x_4, x_5, x_6]^T\), of which \([x_1, x_2]^T\) is the position vector; \([x_3, x_4]^T\) is the velocity vector; \([x_5, x_6]^T\) is the acceleration vector; and \(\omega_1(t)\) and \(\omega_2(t)\) are the independent jerk noise along the \(X\)-axis and \(Y\)-axis, respectively. The initial conditions are \(x_0 = [-30, 0, 1, 0.5, 0.2, 0.1]^T\) and \(\Pi_0 = I_6\).

Typical values of the parameter \(1/\alpha\) for an evasive maneuver are 10–20 s, as suggested in [26]. In the simulations, we choose \(\alpha_1 = 0.1\) and \(\alpha_2 = 0.05\). In addition, the instantaneous variances \(\sigma_{m_1}^2\) and \(\sigma_{m_2}^2\) are set to be

\[
\sigma_{m_1}^2 = \frac{(\sqrt{0.27})^2}{3}[1 + 4 \times 0.2 - 0.3] = 0.135 \text{ m}^2/\text{s}^4
\]

\[
\sigma_{m_2}^2 = \frac{(\sqrt{0.54})^2}{3}[1 + 4 \times 0.2 - 0.3] = 0.27 \text{ m}^2/\text{s}^4
\]

so that \(Q = 0.027I_2\).

We use a sensor network of \(N = 20\) sensors consisting of 12 type-I sensors \(I = \{1, 2, \ldots, 12\}\) and 8 type-II sensors \(I^c = \{13, 14, \ldots, 20\}\), as shown in Fig. 2, to track the maneuvering target. This network is obtained by distributing the sensors randomly over a squared area of 100-m side length. In addition, any two sensors can communicate if the distance between them is smaller than \(r = 30\) m. Moreover, each type-I sensor can observe the distorted position of the target according to the linear model (2) with

\[
C_i = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \end{bmatrix}, R_i = 0.25\sqrt{d_{ij}}I_2 \quad \forall i \in I
\]

where \(d_{ij}\) is the distance between sensor \(i\) and the target. The 513 factor \(\sqrt{d_{ij}}\) in \(R_i\) means that the farther it is from the target, the less information can be observed by the sensor. The other 515 parameters used in the simulations are \(H_i = I_6\), \(\eta = 2\), and \(\alpha_{ij} = \sqrt{1/(1 + d_{ij}^2)}\). Each result presented here is the average 517 of 20 independent runs, except otherwise stated.

Fig. 3 shows the true and estimated trajectories of the target 519 at type-I sensor 2 in one run, from which we can observe that the 520 estimates are close to the true trajectory. This implies that the 521 proposed DOCF algorithm is able to track the target for sensor 522 2. A common sanity test of the estimates is the consistency 523 testing [1], which is crucial for the optimality evaluation. Fig. 3 524 also demonstrates that the proposed DOCF algorithm produces 525 consistent estimates, in other words, the true target positions 526 are almost within the 3\(\sigma\) uncertainty bounds centered at the corresponding target positions.

The simulation results show that up to 528 91.4\% of \(x_j\) fall within the interval \([\hat{x}_{2,j} - 3\sqrt{P_2(j,j)}, \hat{x}_{2,j} + 3\sqrt{P_2(j,j)}]\), \(j = 1, 2\). To qualitatively evaluate the performance of the proposed 532 DOCF algorithm, we introduce the two metrics, namely, the 533
this suggests that there might be an optimal number of type-I sensors, which is one of our future works to find this optimal value.

Finally, we compare the performance of the proposed DOCF algorithm with the centralized Kalman filter, KCF ([8, Algorithm 3]), diffKF ([5, Algorithm 2]), and the nonoptimal algorithms. For a fair comparison, we consider the 571 local measurements based DOCF algorithm denoted by 572 DOCF-M, which is given by

$$\dot{x_i}(t) = (A - G_{i,M}(t)C_{i,M}) \hat{x_i}(t) + G_{i,M}(t) \hat{y}_{i,M}(t) + B_i(t) \sum_{j \in N_i} [\hat{x}_j(t) - \hat{x}_i(t)]$$

where $y_{i,M} = [y_{i1}^T, \ldots, y_{i|N_i|}^T]^T$ is the stacked vector of its 578 neighbors’ measurements and $C_{i,M} = [C_{i,1}^T, \ldots, C_{i,N_i}^T]^T$. It 579 is straightforward to formally derive the optimal consensus 580 filter in this case following the same line as in the previous 581 sections; thus, we omit the details. For the nonoptimal 582 consensus filter, we set $G_{i}(t) = P_{i}(t)C_{i}^TR_{i}^{-1} + 0.01I$ and 583 $\gamma_{ij} = 1/|N_i| \forall j \in N_i$.

Since the centralized Kalman filter, KCF, and diffKF algorithms are inherently proposed for homogeneous sensor networks, here, the simulations are performed over the sensor network consisting of only type-I sensors with the same topology as shown in Fig. 2. As for the DOCF-M algorithm, we 589 consider two cases, namely, $|I^c| = 0$ and $|I^c| = 8$. In Fig. 6, 590 we plot the comparison results with regard to MSD. Clearly, it 591 shows that DOCF-M in the case of $|I^c| = 0$ outperforms the 592 other distributed algorithms. Even if $|I^c| = 8$ type-II sensors 593 are present in the network, MSD still remains at a satisfactory 594 level, which is lower than that of the KCF algorithm. Moreover, 595 over, the result demonstrates that the optimal filter DOCF-M 596 possesses improved estimation accuracy MSD compared with 597 the nonoptimal filter.

Additionally, we define an improvement factor (IF) of the 599 DOCF-M algorithm, as compared with KCF, diffKF, and the 600 nonoptimal algorithms, as

$$IF = \frac{\text{RMSE of the compared one}}{\text{RMSE of DOCF-M}}$$
and optimization problems. Furthermore, we have investigated its convergence property. The theoretical analysis has been validated by simulation results that the estimation errors are exponentially bounded. The simulation results also suggest that the heterogeneous sensor network might be a more appropriate choice for the target tracking problem than the homogeneous one.

Some possible directions remain to be further explored, such as the impact of network topology on the accuracy of tracking.

### VI. Conclusion and Future Work

We have addressed the distributed tracking problem of a maneuvering target over heterogeneous sensor networks. A novel DOCF is proposed to take the heterogeneity of sensor ability into account by solving the optimal control problems.

### APPENDIX A

#### PROOF OF LEMMA 1

**Necessity:** Subtracting (1) from (3) and using (4), one obtains for \( i \in I \)

\[
\hat{e}_i(t) = F_i(t)\hat{e}_i(t) + (F_i(t) + G_i(t)C_i - A)x(t) + G_i(t)v_i(t) + H_i(t)\sum_{j \in N_i} a_{ij} [e_j(t) - c_i(t)] - Bw(t) \tag{36}
\]

and for \( i \in I^c \)

\[
e_i(t) = \sum_{j \in N_i} \gamma_{ij}(t)e_j(t) + \sum_{j \in N_i} \gamma_{ij}(t) - 1]x(t) \tag{37}
\]

The requirement of unbiasedness means that \( \mathbb{E}\{e_i(t)\} = 0 \) for all \( i \in V \). Noting that \( \mathbb{E}\{v_i(t)\} = 0 \), we obtain for both sides of (36) and (37), we have for \( i \in V \) and \( \mathbb{E}\{u(t)\} = 0 \).

Thus, by taking expectation on both sides of (36) and (37), we have for \( i \in V \)

\[
\mathbb{E}\{e_i(t)\} = 0 \quad \forall i \in V \Rightarrow \sum_{j \in N_i} \gamma_{ij}(t) = 1 \Rightarrow \mathbb{E}\{x(t)\} = 0.
\]

### APPENDIX B

#### PROOF OF LEMMA 2

For \( i \in V \), we have

\[
\hat{e}_i(t) = F_i(t)\hat{e}_i(t) + H_i(t)\sum_{j \in N_i} a_{ij} [e_j(t) - c_i(t)] \tag{38}
\]

where \( \hat{e}_i(t) = \mathbb{E}\{e_i(t)\} \), for all \( i \in V \). Since \( \sum_{j \in N_i} \gamma_{ij}(t) = 1 \), it follows from (37) and Proposition 1[16] that

\[
\hat{e}_i(t) = \sum_{j \in I} \gamma_{ij}(t)e_j(t) \tag{39}
\]

where \( \gamma_{ij} \) satisfy \( \sum_{j \in I} \gamma_{ij}(t) = 1 \). Let \( \hat{e}(t) = [\hat{e}_i(t), \hat{e}_{M}(t), 655 \ldots, \hat{e}_{M}(t)]^T \), then, by substituting (39) to (38), one gets the linear system

\[
\hat{e}(t) = \Psi(t)\hat{e}(t), \quad \text{where} \quad \Psi(t) = \text{the coefficient matrix.}
\]

It is known that the solution is given by \( \hat{e}(t) = \Psi(t,0)\hat{e}(0) \), and \( \Psi(t,0) \) is the state transition matrix. By the assumption, we have \( \hat{e}(0) = 0 \). As a result, \( \hat{e}(t) = 0 \), for all \( t > 0 \). It thus follows from (39) that \( \hat{e}_i(t) = 0, \forall i \in I, t > 0 \). Therefore, the 661 estimate \( \hat{e}_i(t) \) is unbiased for all \( i \in V, t > 0 \).
APPENDIX B

DERIVATION OF (10)–(13)

A) Derivation of (10): Under the unbiased requirement, the estimation error $e_i(t)$ satisfies

$$
\dot{e}_i(t) = (A - G_i(t)C_i) e_i(t) + H_i(t) \sum_{j \in N_i} a_{ij} [e_j(t) - e_i(t)] + G_i v_i(t) - B w(t) \quad \forall i \in I. \tag{40}
$$

The solution $e_i$ to (40) can be expressed as

$$
e_i(t) = \Phi_i(t, 0)e_i(0)
$$

$$
+ \int_0^t \Phi_i(t, \tau) [\psi_i(\tau) + G_i(\tau)v_i(\tau) - Bw(\tau)] d\tau
$$

where $\Phi_i(t, \tau)$ is the state transition matrix corresponding to $A - G_i(t)C_i - H_i(t) \sum_{r \in N_i} a_{ir}$ and $\psi_i(\tau) \triangleq H_i(t) \sum_{r \in N_i} a_{ir} e_r(\tau)$. Since $x(0)$, $w(t)$, and $v_i(t)$, $\forall i \in I$, 671 are independent, we have

$$
\mathbb{E} \left\{ e_i(t) v_j^T(t) \right\} = \int_0^t \Phi_i(t, \tau) \mathbb{E} \left\{ \psi_i(\tau) v_j^T(t) \right\} d\tau \tag{S1}
$$

$$
+ \int_0^t \Phi_i(t, \tau) G_i(\tau) \mathbb{E} \left\{ v_i(\tau) v_j^T(t) \right\} d\tau \tag{S2}
$$

$$
\mathbb{E} \left\{ e_i(t) w^T(t) \right\} = \int_0^t \Phi_i(t, \tau) \mathbb{E} \left\{ \psi_i(\tau) w^T(t) \right\} d\tau \tag{S3}
$$

$$
- \int_0^t \Phi_i(t, \tau) B \mathbb{E} \left\{ w(\tau) w^T(t) \right\} d\tau \tag{S4}
$$

Furthermore, we can derive from (40) that

$$
\hat{P}_{ij}(t) = \mathbb{E} \left\{ \dot{e}_i(t) \dot{e}_j^T(t) \right\} + \mathbb{E} \left\{ e_i(t) \dot{e}_j^T(t) \right\}
$$

$$
= (A - G_i(t)C_i) P_{ij}(t) + P_{ij}(t) (A - G_j(t)C_j)^T
$$

$$
+ H_i(t) P_{N_i,j}(t) + P_{N_i,j} H_j^T(t)
$$

$$
+ G_i(t) \mathbb{E} \left\{ v_i(t) e_j^T(t) \right\} + \mathbb{E} \left\{ e_i(t) v_j^T(t) \right\} G_j^T(t)
$$

$$
- B \mathbb{E} \left\{ w(t) e_j^T(t) \right\} - \mathbb{E} \left\{ e_i(t) w^T(t) \right\} B^T
$$

which, together with the preceding analysis, yields (10).

Derivation of (11) and (12): We only prove (11) since (12) can be deduced in a similar manner. The unbiasedness reduces 684 (37) to be

$$
e_i(t) = \sum_{j \in I^c} \tilde{\gamma}_{ij}(t) e_j(t) \quad \forall i \in I^c \tag{41}
$$

which implies that, for any two $i \in I$ and $j \in I^c$, one can write 686 $\hat{P}_{ij}$ as follows:

$$
\hat{P}_{ij}(t) = \mathbb{E} \left\{ \dot{e}_i(t) e_j^T(t) \right\} + \sum_{l \in I^c} \tilde{\gamma}_{jl} \mathbb{E} \left\{ e_i(t) \dot{e}_l^T(t) \right\}.
$$

Then, repeating the similar arguments as in the derivation of 688 (10) and bearing in mind the constraint (7), we can obtain (11).

B) Derivation of (13): In view of (41), $P_{ij}$, $\forall i, j \in I^c$, can be expressed as

$$
P_{ij}(t) = \mathbb{E} \left\{ \sum_{h, l \in I^c} \tilde{\gamma}_{ih} e_h(t) \sum_{t \in I^c} \tilde{\gamma}_{jl} e_l^T(t) \right\}
$$

$$
= \sum_{h, l \in I^c} \tilde{\gamma}_{ih} \tilde{\gamma}_{jl} P_{hl}(t).
$$

which is just (13).

APPENDIX C

PROOF OF LEMMA 2

In order to compute the optimal gain matrix $G_i$ corresponding to the cost function $J_i(P_i)$, we first note that this 696 minimization problem under the constraint (14) is analogous to the classic optimal control problem, where, now, $P_i$ can be considered as the state of a system and $G_i$ as the control input. Therefore, the minimum principle of Pontryagin can be used to check here as in [28], where it was employed to derive the centralized 701 Kalman–Bucy filter.

To do this, let $t_f$ be the terminal time and the cost function becomes $J_i(P_i) = \text{tr} \left[ \tilde{P}_i(t_f) \right]$, $i \in I$. Then, the optimal control problem with free final state $\tilde{P}_i(t_f)$ is readily followed by introducing the Hamiltonian function $H_i(\tilde{P}_i, G_i, \Sigma_i, t) = \text{tr} \left[ \tilde{P}_i(t) \Sigma_i^T(t) \right]$, where $\Sigma_i$ is an $n \times n$ matrix of Lagrange multipliers.

According to the minimum principle of Pontryagin [29], the optimal matrix $G_i^*(t)$ and the corresponding matrix $\Sigma_i^*(t)$ must 700
711 satisfy the following conditions:

\[-\Sigma_i^*(t) = \frac{\partial H_i}{\partial P_i} \left( \dot{P}_i^*(t), G_i^*(t), \Sigma_i^*(t), t \right)\]

\[0 = \frac{\partial H_i}{\partial G_i} \left( \dot{P}_i^*(t), G_i^*(t), \Sigma_i^*(t), t \right)\]

\[\Sigma_i^*(t_f) = \frac{\partial J_i}{\partial P_i} \left( \dot{P}_i^*(t_f) \right)\].

712 In view of matrix calculus, it can be derived that

\[\frac{\partial H_i}{\partial P_i} = \left( A - G_i^*(t)C_i - \sum_{r \in N_i} a_{ir} H_i(t) \right)^T \Sigma_i^*(t)\]

\[+ \Sigma_i^*(t) \left( A - G_i^*(t)C_i - \sum_{r \in N_i} a_{ir} H_i(t) \right)\]

\[= \Sigma_i^*(t) \left( A - G_i^*(t)C_i - \sum_{r \in N_i} a_{ir} H_i(t) \right)\]

\[\frac{\partial H_i}{\partial G_i} = \Sigma_i^*(t) \left( A - G_i^*(t)C_i - \sum_{r \in N_i} a_{ir} H_i(t) \right)\]

\[\frac{\partial J_i}{\partial P_i} = I_n.\]

713 From the preceding equations, one can obtain

\[\dot{\Sigma}_i^*(t) = -\left( A - G_i^*(t)C_i - \sum_{r \in N_i} a_{ir} H_i(t) \right)^T \Sigma_i^*(t)\]

\[-\Sigma_i^*(t) \left( A - G_i^*(t)C_i - \sum_{r \in N_i} a_{ir} H_i(t) \right)\]

714 with the terminal condition \(\Sigma_i^*(t_f) = I_n\). It then follows from 715, Proposition 1.1[30] that the matrix \(\Sigma_i^*(t) > 0, \forall t \geq 0\), and 716 thus nonsingular. In consequence, (43) together with (46) gives

\[G_i^*(t) R_i(t) = P_i^*(t) C_i^T, \forall t \geq 0.\]

Since \(R_i(t)\) is assumed to be 718 positive definite \(R_i(t) > 0\), it implies that the optimal gain 719 matrix \(G_i^*\) is given by \(G_i^*(t) = P_i^*(t) C_i^T R_i^{-1}(t)\).

720 Appendix D

Proof of Lemma 3

722 Note that \(\tilde{g}_{ij}\) are constrained by (7). By setting \(j = i\) in (13'), 723 one can collect (13') and (7) into the following optimization 724 problem to determine the optimal \(\tilde{g}_{ij}\):

\[
\begin{align*}
\min_{\tilde{g}_{ij; k=1,2,\ldots,p_i}} & \sum_{h=1}^{p_i} \sum_{l=1}^{p_i} \tilde{g}_{ijh} \tilde{g}_{ijl} \text{tr}[\hat{P}_{jnh}] \\
\text{s.t.} & \sum_{k=1}^{p_i} \tilde{g}_{ijh} = 1 \\
& \tilde{g}_{ijh} > 0, \quad k = 1, 2, \ldots, p_i.
\end{align*}
\]

Substituting \(\tilde{g}_{ij} = 1 - \sum_{h=2}^{p_i} \tilde{g}_{ijh}\) into the objective function 725 of (48) results in

\[
\begin{align*}
\min_{\tilde{g}_{ij; k=2,3,\ldots,p_i}} & \sum_{h=2}^{p_i} \sum_{k=2}^{p_i} \tilde{g}_{ijh} \tilde{g}_{ijl} \text{tr}[\hat{P}_{jnh}] \\
& + \left( 1 - \sum_{h=2}^{p_i} \tilde{g}_{ijh} \right) \sum_{l=2}^{p_i} \tilde{g}_{ijl} \text{tr}[\hat{P}_{jil}] \\
& + \left( 1 - \sum_{l=2}^{p_i} \tilde{g}_{ijl} \right) \text{tr}[\hat{P}_{jil}] \\
\text{s.t.} & \tilde{g}_{ijh} > 0, \quad k = 2, 3, \ldots, p_i.
\end{align*}
\]

The preceding minimization problem can be solved by the 727 method of Lagrangian multipliers. Writing the Kuhn–Tucker 728 conditions, we have

\[
\sum_{l=2}^{p_i} \tilde{g}_{ijh} \text{tr}[\hat{P}_{jnh}] = \left( 1 - \sum_{l=2}^{p_i} \tilde{g}_{ijl} \right) \text{tr}[\hat{P}_{jnh}] = 0
\]

and \(\tilde{g}_{ijh} > 0\), for all \(k = 2, 3, \ldots, p_i\). Then, the compact form 730 can be expressed as

\[
U_i \gamma_i^* = V_i, \quad \text{and} \quad \tilde{g}_{ij} > 0
\]

where \(\gamma_i^* \triangleq [\tilde{g}_{ij1}, \tilde{g}_{ij2}, \ldots, \tilde{g}_{ijp_i}]^T\). This means that the optimal 732 \(\tilde{g}_{ij}\) can be obtained by solving the linear equation (49). The 733 lemma thus follows.

Appendix E
Proof of Lemma 4

Note that \(\tilde{g}_{ij}\) satisfies the condition (7). From (41), we obtain 737 for all \(i \in \mathcal{I}\)

\[
\sum_{j \in N_i} a_{ij} (e_j - e_i)
\]

\[
= \sum_{j \in N_i \setminus \mathcal{I}_i} a_{ij} (e_j - e_i) + \sum_{j \in \mathcal{I}_i} a_{ij} \tilde{g}_{ijk} (e_k - e_i)
\]

\[
= \sum_{j \in N_i \setminus \mathcal{I}_i} a_{ij} (e_j - e_i) + \sum_{j \in \mathcal{I}_i \setminus \mathcal{I}} a_{ij} \tilde{g}_{ijk} (e_j - e_i)
\]

\[
= - (I_i^T \otimes I_n) e
\]

where \(\otimes\) is the Kronecker product, and \(I_i \triangleq [l_{i1}, l_{i2}, \ldots, l_{ih}]^T\). 739 Consequently, we can collect (22) and (41) into the compact 740 vector form given in (23). Similar to the derivation of (23), it 741 is not difficult to derive from (10') to (13') and (7) that the 742 dynamics of \(\hat{P}\) can be expressed as (24). This completes the 743 proof.

Acknowledgment

The authors would like to thank the anonymous referees, the 746 Associate Editor, and the Editor-in-Chief for their valuable and 747 constructive comments.
REFERENCES


Shanying Zhu (S’11) received the B.Sc. degree in information and computing science from North China University of Water Resources and Electric Power, Zhengzhou, China, in 2006 and the M.Sc. degree in applied mathematics from Huazhong University of Science and Technology, Wuhan, China, in 2008. He is currently working toward the Ph.D. degree in the Department of Automation, Shanghai Jiao Tong University, Shanghai, China. His research interests focus on multiagent systems and wireless sensor networks, particularly in coordinate distribution control of mobile robots and distributed detection and estimation in sensor networks and their applications in industrial networks.

Cailian Chen (S’03–M’06) received the B.Eng. and Ph.D. degrees in control automatic control from Yanshan University, Qinhuangdao, China, in 2000 and 2002, respectively, and the Ph.D. degree in control and systems from City University of Hong Kong, Hong Kong, China, in 2006. In 2008, she joined the Department of Automation, Shanghai Jiao Tong University, Shanghai, China, as an Associate Professor. She has authored and/or coauthored a research monograph and over 80 referred international journal and conference papers.

In 2008, she joined the Department of Automation, Shanghai Jiao Tong University, Shanghai, China, as an Associate Professor. She has authored and/or coauthored a research monograph and over 80 referred international journal and conference papers.

Dr. Chen serves as an Associate Editor of ISRN Sensor Networks and an IPC or TPC member of many conferences, including FUZZ-IEEE, IEEE WCIF, IEEE Globecom, IEEE WISARN, IEEE MSC. She was one of the First Prize Winners of a Natural Science Award from the Ministry of Education of China in 2007. She was a recipient of the ‘IEEE Transactions on Fuzzy Systems Outstanding Paper Award’ in 2008. She was honored as ‘Shanghai Pujiang Scholar,’ ‘Chenguang Scholar,’ and ‘SMC Outstanding Young Staff’ of Shanghai Jiao Tong University’ in 2009.

Wireless Networking and Cooperative Control.

Wenshuang Li (S’XX) received the B.S. degree in measurement control technology and instruments from Yanshan University, Qinhuangdao, China, in 2010. She is currently working toward the M.S. degree in control science and engineering at Shanghai Jiao Tong University, Shanghai, China. Her fields of study focus on event-triggered transmission scheme and distributed consensus estimation over wireless sensor networks and their applications in industrial networks.

W. Li is a member of the Center for Intelligent Systems and Wireless Networking and Cooperative Control.
Bo Yang (M’XX) received the Ph.D. degree in electrical engineering from the City University of Hong Kong, Kowloon, Hong Kong, in 2009. He was a Postdoctoral Researcher with the Royal Institute of Technology (KTH), Stockholm, Sweden, from 2009 to 2010 and a Visiting Scholar with the Polytechnic Institute of New York University in 2007. In 2010, he joined Shanghai Jiao Tong University, Shanghai, China, where he was first an Assistant Professor and is currently an Associate Professor. He is also a Cyber Scholar with the Cyber Joint Innovation Center founded by Zhejiang University, Hangzhou, China; Tsinghua University, Beijing, China; and Shanghai Jiao Tong University. His current research interests include game theoretical analysis and nonlinear optimization of communication networks and smart grid.

Dr. Yang is on the Editorial Board of *Digital Signal Processing* (Elsevier) and in the TPC of several international conferences. He has been the principle/co-investigator in several research projects funded by the National Science Foundation of China, the Swedish Governmental Agency for Innovation Systems, and the U.S. Air Force. He is also a member of the Association for Computing Machinery.

Xinping Guan (M’02–SM’04) received the Ph.D. degree in control and systems from Harbin Institute of Technology, Harbin, China, in 2001. In 2007, he joined the Department of Automation, Shanghai Jiao Tong University, Shanghai, China, where he is currently a Distinguished University Professor and the Director of the Key Laboratory of 911 Systems Control and Information Processing, Ministry of Education of China. He has authored and/or coauthored 2 research monographs, more than 120 SCI indexed papers in IEEE Transactions and other well-known international journals, and numerous conference papers. As a Principal Investigator, he has finished/been working on many national key projects. He is the leader of the prestigious Innovative Research Team of the National Natural Science Foundation of China (NSFC) in 2012. His current research interests include wireless sensor networks, ground–air communication of aircrafts, and cognitive radio networks and their applications in industry.

Dr. Guan is a Committee Member of the Chinese Automation Association Council and the Chinese Artificial Intelligence Association Council. He serves as an Associate Editor of the *IEEE Transactions on Systems, Man and Cybernetics—Part C*; an Editorial Board Committee Member for 924 several Chinese journals; and an International Technical Committee Member 926 for a lot of conferences. He was one of the First Prize Winners of a University Natural Science Award from the Ministry of Education of China in 2003 and 2007, respectively. He was a recipient of the “IEEE Transactions on Fuzzy Systems Outstanding Paper Award” in 2008, the Second Prize of the National Natural Science Award of China, and a “National Outstanding Youth Award” from NSFC. He was appointed as “Changjiang Scholar” by the Ministry of Education of China and “State-level Scholar” of “New Century Bai Qianwan Talent Program” of China.
AUTHOR QUERIES

AUTHOR PLEASE ANSWER ALL QUERIES

A mandatory over length page charge of $175 is required for each page in excess of 12 pages for a regular paper and 6 pages for technical correspondences. There is now a MAXIMUM cap of 2 over length pages.

AQ1 = Information with regard to the prior presentation of the paper was changed to: A preliminary version of this paper was presented at the 8th Asian Control Conference at The Splendor Kaohsiung, Kaohsiung, Taiwan, May 2011. Please check if appropriate. Otherwise, please make the necessary changes.

AQ2 = KCF was considered as the abbreviation for “Kalman-consensus filter” and was thus introduced here. Please check if appropriate. Otherwise, please provide the corresponding expanded form.

AQ3 = The portion “Combining the fact that $W > 0$ is positive definite and relation (28) finally enables (32) to be” was changed to “Combining relation (28) and the fact that $W > 0$ is positive definite finally enables (32) to be...” Please check if appropriate. Otherwise, please make the necessary changes.

AQ4 = Please provide the expanded form of ISRN.
AQ5 = Please provide the expanded forms of IPC and TPC.
AQ6 = Please provide the expanded forms of WCCI and MSC.
AQ7 = Please provide IEEE membership history.
AQ8 = Please provide IEEE membership history.
AQ9 = ACM was expanded as “Association for Computing Machinery.” Please check if appropriate. Otherwise, please provide the corresponding expanded form.
AQ10 = Please provide the expanded form of SCI.

END OF ALL QUERIES