ECO-DKF: Event-Triggered and Certifiable Optimal Distributed Kalman Filter under Unknown Correlations

Eduardo Sebastián*, Eduardo Montijano and Carlos Sagüés

Abstract—This paper presents ECO-DKF, the first *E*vent-Triggered and *C*ertifiable *O*ptimal *D*istributed *K*alman *F*ilter. Our algorithm addresses two major issues inherent to Distributed Kalman Filters: (i) fully distributed and scalable optimal estimation and (ii) reduction of the communication bandwidth usage. The first requires to solve an NP-hard optimisation problem, forcing relaxations that lose optimality guarantees over the original problem. Using only information from one-hop neighbours, we propose a tight Semi-Definite Programming relaxation that allows to certify locally and online if the relaxed solution is the optimum of the original NP-hard problem. In that case, ECO-DKF is optimal in the square error sense under scalability and event-triggered one-hop communications restrictions. Additionally, ECO-DKF is a globally asymptotically stable estimator. To address the second issue, we propose an event-triggered scheme from the relaxed optimisation output. The consequence is a broadcasting-based algorithm that saves communication bandwidth, avoids individual communication links and multiple information exchanges within instants, and preserves the optimality and stability properties of the filter.

Index Terms—Certifiability, distributed systems, event-triggered systems, Kalman Filter.

I. INTRODUCTION

The Kalman Filter (KF) [1] is a cornerstone in control theory. Its elegance and optimality has motivated its extension towards distributed setups. However, two major issues prevent a direct deployment in real setups. First, the conservation of the KF optimality. Second, the efficient usage of the communication bandwidth.

Regarding optimality, the works by Olfati-Saber [2], [3] opened an era of consensus-based Distributed Kalman Filters (DKF). However, it is proved that the computation of the estimator's gains is not scalable [4], [5]. Since then, several works have aroused searching for an Optimal DKF (O-DKF) [6]–[8] by either keeping track of the whole network error covariance matrix or neglecting the information encoded in the covariance error matrices. In a recent work [9] it is discussed that, to reach optimal centralised KF performance, the state-of-the-art DKFs typically resort to multiple consensus steps among instants. The absence of solutions which optimally and *scalably* integrate the neighbouring estimates and error covariances motivated the development of Diffusion KFs (DfKF). The most popular approach is the Covariance Intersection (CI) method [10]–[12]. The CI method is generally suboptimal [13], thus arising a necessity of certification. The solution to approach optimality goes through convex optimisation problems [14], sequential algorithms [15] or diffusion steps that require global topology knowledge [16]. Inspired by the DfKF solutions, ECO-DKF gathers the benefits of all the aforementioned solutions and adds the *certifiability* property, which comes from a reformulation of the CI method based on the outer Löwner-John ellipsoid intersection method [17] and a Semi-Definite Programming

(SDP) relaxation. By proving the equivalence between ECO-DKF and the optimal consensus over estimates we can demonstrate global asymptotic stability, independent on any hand-tuned gain.

Regarding communications, the amount of information sent through the network is related to its energy consumption and the available bandwidth [18], so an Event-Triggered (ET) estimator is desirable. The most popular ET rule is the Send-on-Delta (SoD) [19], which computes the difference between estimates [20]–[23], measurements [19], [22], [23] and/or innovations [24]–[26] to trigger the information when the difference exceeds some tuned threshold. SoD is popular because it is easy to compute and agnostic to the employed estimator. In contrast, ECO-DKF exploits the estimator's structure to derive an ET rule which is even easier to compute and does not require tuning. A work conducted by Trimpe and Campi [27] compared the most popular ET rules, concluding that leveraging the error covariance in the decision improves performance, e.g., the Kullback-Leibler (KL) divergence. Regarding KFs, ET Centralised Kalman Filters (CKF) [18], [24], [28]–[30] are subject to a critical arrival rate beyond which the error covariance becomes unbounded [31], underlining the importance of a carefully designed ET protocol. On the other hand, both the SoD [22], [23], [26] and KL [26], [32] protocols are very popular in non-optimal DKFs. In contrast, ECO-DKF is event-triggered-by-construction. This allows us to define an ET rule which is broadcasting-based, inexpensive to compute and avoids individual communication links and multiple information exchanges within instants.

Exploiting these ideas, our main contribution is ECO-DKF, the *first DKF with certifiability guarantees and event-triggered by construction*. If the certification is positive, then ECO-DKF is *optimal in the square error sense* under unknown correlations and potential one-hop event-triggered communications. The proposed algorithm is also *globally asymptotically stable* under mild periodic joint connectivity assumptions. Besides, nodes can *check optimality locally and in real time*. In particular, ECO-DKF exploits the optimisation to derive a broadcasting-based rule, yielding to an algorithm which is *event-triggered by construction*. ECO-DKF is *fully distributed*, works with *heterogeneous* sensor models and does not require *parameter tuning*.

This paper is an extension of [33]. Compared to it, we first improve ECO-DKF from being time-triggered to event-triggered. This leads to new stability and optimality results that hold for mild connectivity settings. Besides, we relax the observability conditions to that of network-observability, a weaker assumption. Consistency of ECO-DKF is proved for the first time, while Lemmas that were presented in [33] are now proved, including additional corollaries. The ET scheme is also a novel contribution, along with the corresponding convergence and optimality guarantees associated to it. Finally, the empirical evaluation of ECO-DKF is improved by adding more time-triggered state-of-the-art DKFs and including new experiments that validate the event-triggered properties of ECO-DKF.

II. PROBLEM FORMULATION

The target system is described by linear dynamics

$$\mathbf{x}(k+1) = \mathbf{A}\mathbf{x}(k) + \mathbf{w}(k), \tag{1}$$

E. Sebastián, E. Montijano and C. Sagüés are associated with the Instituto de Investigación en Ingeniería de Aragón, Universidad de Zaragoza, Spain {esebastian, emonti, csagues}@unizar.es. *Corresponding author.

where $k \in \mathbb{N}_{\geq 0}$ denotes the discrete time, $\mathbf{x} \in \mathbb{R}^n$ is the state of the system to estimate, $\mathbf{A} \in \mathbb{R}^{n \times n}$ is a matrix comprising the dynamics of the target system, and $\mathbf{w}(k) \sim \mathcal{N}(\mathbf{0}, \mathbf{Q})$ is a white process. The system is tracked by a network of sensors described by a directed communication graph $G(k) = (V, E(k))$. The number of nodes is $N = |V|$, where $|\cdot|$ denotes cardinality. $\mathcal{N}_i(k) = \{j | (i,j) \in E(k)\}$ is the set of neighbours of node i at instant k and $\mathcal{J}_i(k) = \mathcal{N}_i(k) \cup i$. We use $\lfloor \cdot \rfloor_{ij}$ to denote the ij-th element or block component of a matrix. Sensors are described by a linear model

$$\mathbf{z}_i(k) = \mathbf{H}_i \mathbf{x}(k) + \mathbf{v}_i(k). \tag{2}$$

In this expression, $\mathbf{z}_i \in \mathbb{R}^{m_i}$ is the measurement of node i, $\mathbf{H}_i \in \mathbb{R}^{m_i \times n}$ is the (unbiased) sensor model of node i, m_i is the dimension of $\mathbf{z}_i$ and $\mathbf{v}_i(k) \sim \mathcal{N}(\mathbf{0}, \mathbf{R}_i)$ is a white process. The measurements are independent between nodes. Notice that we are considering heterogeneous sensors. Unless unclear, from now on we will omit the dependencies with k. Regarding observability, we assume *network-observability*, defined as follows.

Definition 1. *System* (1) *is network-observable if the pair* $(\mathbf{A}, [\mathbf{H}_1^T, \mathbf{H}_2^T, \dots, \mathbf{H}_N^T]^T)$ *is observable.*

The objective of the network is to cooperatively estimate the state of system (1). Each node has an estimate of the state of system (1), $\hat{\mathbf{x}}_i$, with associated error covariance matrix $\hat{\mathbf{P}}_i := \mathrm{E}[(\hat{\mathbf{x}}_i - \mathbf{x})(\hat{\mathbf{x}}_i - \mathbf{x})^T]$. Besides, similarly to other typical filter algorithms, the nodes use an auxiliary variable for the prediction stage, $\bar{\mathbf{x}}_i$, and $\bar{\mathbf{P}}_i := \mathrm{E}[(\bar{\mathbf{x}}_i - \mathbf{x})(\bar{\mathbf{x}}_i - \mathbf{x})^T]$. In the classical CKF, the estimate $\hat{\mathbf{x}}_i(k-1), \hat{\mathbf{P}}_i(k-1)$ is first used to obtain the prediction $\bar{\mathbf{x}}_i(k), \bar{\mathbf{P}}_i(k)$ by propagation through the target dynamics. Then, the measurement is used to correct the prediction, leading to the updated estimate $\hat{\mathbf{x}}_i(k), \hat{\mathbf{P}}_i(k)$. The CKF is not scalable in networked applications because a central computation unit is needed to aggregate the measurements from all the nodes. An alternative is to develop a DKF, where each node uses only local and neighbouring information.

Given, $\bar{\mathbf{P}}_i(k)$, we define $\bar{\mathbf{S}}_i(k) = \bar{\mathbf{P}}_i^{-1}(k)$ as the predicted error covariance matrices in information form. In this paper we are interested in viewing these matrices as ellipsoids, defined as follows.

Definition 2. *Given* $\bar{\mathbf{S}}_i$ *and assuming unbiased sensors, the ellipsoid* ε_i^i *is* $\varepsilon_i^i := \{ \mathbf{x} \,|\, \mathbf{x}^T \bar{\mathbf{S}}_i^{-1} \mathbf{x} \leq 1 \}$.

In our proposal, this absence of global knowledge leads to an NP-hard optimisation problem, caused by the unknown correlations between the nodes' estimates, forcing the use of convex relaxations to solve it. Despite enabling tractability, the solution of the relaxed problem is not guaranteed to be the optimum of the original problem. This motivates the need of certifiability on the optimisation, formally defined as follows:

Definition 3 (**From Definition 19 in [34]**). *Given an optimization problem* $\mathbb{O}(\mathbb{D})$ *that depends on input data* $\mathbb{D}$, *we say that an algorithm* $\mathbb{A}$ *is certifiable if, after solving* $\mathbb{O}(\mathbb{D})$, $\mathbb{A}$ *either provides a certificate for the optimality of its solution or declares failure otherwise.*

Besides, to avoid divergence of the estimation, the result of the optimisation must be consistent, defined as follows:

Definition 4. *An approximation* $\tilde{\mathbf{P}}$ *of* $\mathbf{P}$ *and* $\tilde{\mathbf{x}}$ *of* $\mathbf{x}$ *is consistent if and only if* $\tilde{\mathbf{P}} - \mathbf{P} \succeq \mathbf{0}$ *and* $\mathrm{E}[(\tilde{\mathbf{x}} - \mathbf{x})] = \mathbf{0}$, *where* $\succeq$ *means positive semi-definiteness.*

Many DKFs rely on time-triggered communications of all the nodes, requiring an intensive bandwidth usage. Instead, event-triggered communications save energy and bandwidth. In this regard, we define two concepts. The first is Periodic Joint Connectivity (PJC):

Definition 5. *Consider a sequence of graphs* $\{G(k), G(k + i), \dots, G(k + t)\}$ *with* $t \in \mathbb{N}$. *If* $\mathbf{G}(k,t) = \cup_{i=0}^{t} G(k+i)$ *is a strongly connected graph, then* $\mathbf{G}(k,t)$ *is a periodic joint connected graph with period* t *at instant* k.

Second, we define a metric to measure the communication usage in the network. Before, we define the concept of "triggering function":

Definition 6. *A triggering function* $g_i(\cdot, k) \in \{0, 1\}$ *is an indicator function that depends only on local information at node* i *and such that, if* $g_i(\cdot, k) = 1$, *then node* i *broadcasts at instant* k, *and, if* $g_i(\cdot, k) = 0$, *then node* i *does not broadcast at instant* k.

Definition 7. *The Number of Broadcasts (*NoB*) at instant* k *in network* $G(k)$ *is*

$$\mathrm{NoB}(k) := \sum_{i=1}^{N} g_i(\cdot, k). \tag{3}$$

We now formulate the problem addressed in the paper.

Problem 1. *Find an event-triggered stable algorithm that certificates locally and in real time if, at each* k, *each node* i *minimises* $\mathrm{E}[||\hat{\mathbf{x}}_i(k) - \mathbf{x}(k)||^2]$, *under the following restrictions:*

1) *Locality: At instant* k *and* $\forall i$, *node* i *only uses* $\mathbf{A}$, $\mathbf{H}_i$, $\mathbf{Q}$, *and* $\mathbf{R}_i$ *as parameters. From instant* k *to* $k+1$ *and* $\forall i$, *node* i *stores* $\hat{\mathbf{x}}_i(k)$, $\hat{\mathbf{P}}_i(k)$, $\mathbf{A}$, $\mathbf{Q}$, $\mathbf{H}_i$ *and* $\mathbf{R}_i$.
2) *One-hop communication: At instant* k *and* $\forall i$, *node* i *communicates at most once with its neighbours* $j \in \mathcal{N}_i(k)$.
3) *Consistency and connectivity: The ET rule ensures consistency* $\forall i, k$, *and PJC of* $G(k)$ $\forall k$ *and some finite* t.
4) *Network-observability*.

Remark 1. *From Problem 1, each node seeks to minimize, at each instant* k, $E[||\hat{\mathbf{x}}_i(k) - \mathbf{x}(k)||^2]$. *Therefore, globally, the network minimises, at each instant* k, $\sum_i E[||\hat{\mathbf{x}}_i(k) - \mathbf{x}(k)||^2]$. *Note that this definition holds because the minimisation refers to the current instant rather than a finite or infinite horizon of time.*

We will use * to denote optimality in the sense of Problem 1. We assume that communication delays, quantisation effects and dropouts are negligible, a common assumption in many works ([24], [25], [28], [35] among others).

III. ECO-DKF ALGORITHM

The proposed solution for Problem 1 is the ECO-DKF algorithm. We first offer an overview of ECO-DKF and then, in Sections IV, V and VI, we formally study the main properties of the estimator.

ECO-DKF algorithm is based on a novel optimization problem to aggregate predicted KF estimates. We propose to use the outer Löwner-John (LJ) method [36], which computes the smallest ellipsoid that surrounds the intersection of a set of ellipsoids,

$$\bar{\mathbf{S}}_i^*, \boldsymbol{\lambda}_i^* = \underset{\bar{\mathbf{S}}, \boldsymbol{\lambda}}{\arg\min} \quad \mathrm{Tr}(\bar{\mathbf{S}}^{-1}) \tag{4a}$$

$$s.t. \quad \mathbf{0} \prec \bar{\mathbf{S}} \preceq \sum_{j=1}^{|\mathcal{J}_i|} \lambda_j \bar{\mathbf{S}}_j, \tag{4b}$$

$$\sum_{j=1}^{|\mathcal{J}_i|} \lambda_j \leq 1, \ \lambda_j \geq 0 \ \forall j \in \mathcal{J}_i, \tag{4c}$$

where $\bar{\mathbf{S}}_j = \bar{\mathbf{P}}_j^{-1}$ characterise the ellipsoids, $\mathrm{Tr}(\cdot)$ is the trace of a matrix, λ_j is the j-th element of vector $\boldsymbol{\lambda}$, and $\prec$ and $\preceq$ denote definiteness and semi-definiteness. The selection of the trace as the optimisation cost function follows from minimising the square error $\mathrm{E}[||\hat{\mathbf{x}}_i - \mathbf{x}||^2]$ in Problem 1, proved later on Theorem 2. The input

of the outer LJ method is the predicted covariance, in information form, from node i and its neighbours. Meanwhile, $\boldsymbol{\lambda}_i$ weights the importance of each prediction. The output of (4) is used to aggregate the predictions as follows

$$\bar{\mathbf{x}}_i^* = \bar{\mathbf{P}}_i^* \sum_{j \in \mathcal{J}_i(k)} \lambda_{ij}^* \bar{\mathbf{s}}_j \text{ and } \bar{\mathbf{P}}_i^* = (\bar{\mathbf{S}}_i^*)^{-1}. \tag{5}$$

The study of this aggregation is developed in Section IV, including the certification guarantees on (4) which permits the formal results about optimality in Section V. To the best of our knowledge, this is the first time optimisation problem (4) is applied in a DKF.

Interestingly, the output of (4) allows to develop a novel ET rule to decide if a node broadcasts a message to its neighbours:

$$g_i(\boldsymbol{\lambda}_i^*(k-1), k) \equiv$$
$$g_i(\boldsymbol{\lambda}_i^*) = \begin{cases} 1, \text{ if } (\max(\boldsymbol{\lambda}_i^*) = \lambda_{ii}^*) \vee (\mathcal{J}_i = \{i\} \wedge \text{Ber}(p) = 1) \\ 0, \text{ otherwise} \end{cases} \tag{6}$$

where $\lambda_{ii}^* \in [0, 1]$ is the weight of $\boldsymbol{\lambda}_i^*$ associated to node i, $\vee$ is the logical "OR", $\wedge$ is the logical "AND", and $\text{Ber}(p)$ represents the Bernoulli distribution of probability p. The messages are defined by $\text{msg} = \{\mathbf{U}_i, \mathbf{u}_i, \bar{\mathbf{S}}_i, \bar{\mathbf{s}}_i\}$, where

$$\mathbf{u}_i = \mathbf{H}_i^T \mathbf{R}_i^{-1} \mathbf{z}_i, \ \mathbf{U}_i = \mathbf{H}_i^T \mathbf{R}_i^{-1} \mathbf{H}_i, \ \bar{\mathbf{s}}_i = \bar{\mathbf{P}}_i^{-1} \bar{\mathbf{x}}_i, \ \bar{\mathbf{S}}_i = \bar{\mathbf{P}}_i^{-1}.$$

The quantities are in information form [37] to comply with restriction 1 of Problem 1. We use the term "broadcast" because, when a node sends its information, there is no distinction to which neighbours the message arrives. A selective communication requires establishment of particular communication channels and a priori knowledge on the set of neighbours. Besides, we do not allow multiple communications within an instant.

Now, we are ready to present ECO-DKF in Algorithm 1. Given $\hat{\mathbf{x}}_i^*(k-1)$, $\hat{\mathbf{P}}_i^*(k-1)$ and $\boldsymbol{\lambda}_i^*(k-1)$, each node first computes the prediction of the state through the target system dynamics,

$$\bar{\mathbf{P}}_i(k) = \mathbf{A}\hat{\mathbf{P}}_i^*(k-1)\mathbf{A}^T + \mathbf{Q}, \tag{7}$$
$$\bar{\mathbf{x}}_i(k) = \mathbf{A}\hat{\mathbf{x}}_i^*(k-1). \tag{8}$$

Then, each node obtains its measurement $\mathbf{z}_i(k)$ and decides whether to broadcast a message or not according to the ET rule (6). Node i receives messages from broadcasting neighbours. The information from the messages is aggregated to compute the values that will be used in the correction step. In the case of the measurements, this is direct because they are independent

$$\mathbf{y}_i = \sum_{j \in \mathcal{J}_i} \mathbf{u}_j \text{ and } \mathbf{Y}_i = \sum_{j \in \mathcal{J}_i} \mathbf{U}_j. \tag{9}$$

On the other hand, the aggregation of predictions is harder because we must consider the unknown correlations between them. To solve this, we use (4) and (5).

Finally, each node calculates the estimated error covariance matrix

$$\hat{\mathbf{P}}_i^* = \left(\bar{\mathbf{S}}_i^* + \mathbf{Y}_i\right)^{-1}. \tag{10}$$

This is analogous to the correction step of the information form [37] DKF presented by Olfati-Saber (Eq. (10) in [4]), where the optimal predicted information covariance matrix, $\bar{\mathbf{S}}_i^*$, is fused with all the measurement information covariance matrices, $\mathbf{Y}_i$, because the measurements are uncorrelated and $\bar{\mathbf{S}}_i^*$ is consistent (proven in Proposition 3). The estimated state is obtained as follows:

$$\hat{\mathbf{x}}_i^* = \bar{\mathbf{x}}_i^* + \hat{\mathbf{P}}_i^*(\mathbf{y}_i - \mathbf{Y}_i\bar{\mathbf{x}}_i^*). \tag{11}$$

It is noteworthy to see that Eq. (11) is similar to

$$\hat{\mathbf{x}}_i = \bar{\mathbf{x}}_i + \hat{\mathbf{P}}_i^* (\mathbf{y}_i - \mathbf{Y}_i\bar{\mathbf{x}}_i) + \gamma\hat{\mathbf{P}}_i^* \sum_{j \in \mathcal{N}_i} (\bar{\mathbf{x}}_j - \bar{\mathbf{x}}_i), \tag{12}$$

which comes from the consensus-based DKF in [3]. However, in ECO-DKF the consensus is implicit in the optimisation, overcoming the dependence for stability on the parameter γ.

We briefly discuss the communication and computational burden. Regarding communication, the size of each message is constant with the number of nodes, so the proposal is scalable. Besides, nodes do not need any global knowledge of the topology nor sensor models of neighbours. Regarding the computational cost, the bottleneck is (4). However, the current hardware can solve large instances of this optimisation problem in real-time [34].

Algorithm 1 ECO-DKF in node i

1: Initialisation: $\hat{\mathbf{P}}_i^*(0) = \mathbf{P}_0$, $\hat{\mathbf{x}}_i^*(0) = \mathbf{x}_0$, $\boldsymbol{\lambda}_i^*(0) = \boldsymbol{\lambda}_0$
2: **while** True **do**
3:     Prediction:
$$\bar{\mathbf{P}}_i(k) = \mathbf{A}\hat{\mathbf{P}}_i^*(k-1)\mathbf{A}^T + \mathbf{Q}$$
$$\bar{\mathbf{x}}_i(k) = \mathbf{A}\hat{\mathbf{x}}_i^*(k-1)$$
4:     Measurement: get $\mathbf{z}_i(k)$
5:     Broadcasting:
6:         **if** $g_i(\boldsymbol{\lambda}_i^*(k-1), k) == 1$ **then**
7:             Send $\{\mathbf{U}_i(k), \mathbf{u}_i(k), \bar{\mathbf{S}}_i(k), \bar{\mathbf{s}}_i(k)\}$
8:         **end if**
9:         Receive $\{\mathbf{U}_j(k), \mathbf{u}_j(k), \bar{\mathbf{S}}_j(k), \bar{\mathbf{s}}_j(k)\} \ \forall j \in \mathcal{N}_i(k)$
10:   Aggregation of measurements' data:
$$\mathbf{Y}_i(k) = \sum_{j \in \mathcal{J}_i(k)} \mathbf{U}_j(k)$$
$$\mathbf{y}_i(k) = \sum_{j \in \mathcal{J}_i(k)} \mathbf{u}_j(k)$$
11:   Aggregation of predictions' data:
$$\bar{\mathbf{S}}_i^*(k), \boldsymbol{\lambda}_i^*(k) \leftarrow \text{Solution of (4)}$$
$$\bar{\mathbf{P}}_i^*(k) = (\bar{\mathbf{S}}_i^*(k))^{-1}$$
$$\bar{\mathbf{x}}_i^*(k) = \bar{\mathbf{P}}_i^*(k) \sum_{j \in \mathcal{J}_i(k)} \lambda_{ij}^*(k)\bar{\mathbf{s}}_j(k)$$
12:   Correction:
$$\hat{\mathbf{P}}_i^*(k) = (\bar{\mathbf{S}}_i^*(k) + \mathbf{Y}_i(k))^{-1}$$
$$\hat{\mathbf{x}}_i^*(k) = \bar{\mathbf{x}}_i^*(k) + \hat{\mathbf{P}}_i^*(k)(\mathbf{y}_i(k) - \mathbf{Y}_i(k)\bar{\mathbf{x}}_i^*(k))$$
13: **end while**

IV. CERTIFIABLE COVARIANCE BOUNDING

To achieve the optimality pursued in Problem 1 it is necessary to deal with the optimal aggregation of neighbouring predictions under unknown correlations. This fusion can be described as finding the minimum volume ellipsoid containing the intersection of the p ellipsoids [17] formed by the matrices $\bar{\mathbf{P}}_j$ in the set $\mathcal{J}_i(k)$. Thus, we first define the concept of intersection of ellipsoids.

Definition 8. *The intersection of p ellipsoids, at node i, is the convex set* $\mathcal{F}_i := \varepsilon_1^i \cap \ldots \cap \varepsilon_j^i \cap \ldots \cap \varepsilon_p^i$.

From Definitions 2 and 8, volume minimisation can be transformed into a maximisation over the information matrices. In particular since the expected square error is equivalent to the trace of $\hat{\mathbf{P}}_i^*$, the objective of ECO-DKF is to optimise this metric.

Problem 2 (Adapted from [17]). *Find* $(\varepsilon_i^i)^*$ *such that it contains* $\mathcal{F}_i$ *and* $\text{Tr}(\bar{\mathbf{S}}_i^*)$ *is maximised.*

Problem 2 is solved by the following optimisation problem [38]:

$$\max_{\mathbf{X}} \quad \text{Tr}(\mathbf{X}) \tag{13a}$$
$$s.t. \quad \mathbf{X} \succeq \mathbf{0}, \ \text{Tr}(\mathbf{X}\bar{\mathbf{S}}_j) \leq 1 \ \forall j \in \mathcal{J}_i(k), \tag{13b}$$
$$\text{rank}(\mathbf{X}) = 1. \tag{13c}$$

This problem is NP-hard due to the non-convex constraint (13c). Thus, it is necessary to find a convex relaxation to make the

optimisation tractable. The simplest solution is to drop the non-convex constraint,

$$\max_{\mathbf{X}} \quad \text{Tr}(\mathbf{X}) \tag{14a}$$

$$s.t. \quad \mathbf{X} \succeq \mathbf{0}, \ \text{Tr}(\mathbf{X}\bar{\mathbf{S}}_j) \leq 1 \ \ \forall j \in \mathcal{J}_i(k), \tag{14b}$$

However, it is not possible to use (14) in ECO-DKF because it does not provide an estimate nor an error covariance matrix, but $\mathbf{X}$. Instead, we propose to use the outer Löwner-John (LJ) method [36], which leads to the relaxed optimisation problem in (4). The input of the outer LJ method is the $|\mathcal{J}_i(k)|$ ellipsoids described by matrices $\bar{\mathbf{S}}_j \ \forall j \in \mathcal{J}_i(k)$. Then, it computes the smallest ellipsoid that includes $\mathcal{F}_i(k)$. The intersection always exist since the ellipsoids are all centered at $\mathbf{0}$ [38]. From (4b), $\bar{\mathbf{S}}_i^*$ is upper-bounded by the weighted sum of matrices $\bar{\mathbf{S}}_j \ \forall j \in \mathcal{J}_i(k)$. Therefore, $\bar{\mathbf{S}}_i^*$ depends on the topology and the target system dynamics, for they determine the set of predicted error covariance matrices considered in the optimisation (Eq. (4b)) and the value of each prediction (Eqs. (7)-(8)) respectively. Problem (4) outputs an estimate and error covariance matrix. Nevertheless, problem (14) is important because it enables the certification of optimality. This is formally stated in the next Proposition, whose proof can be found in [33]:

Proposition 1. *Let $\mathbf{X}^*$ be the solution of* (14)*. Define*

$$\mathbb{C}_i := \text{rank}\left(\mathbf{X}^*\right) \ \text{ and } \rho_i := \text{Tr}(\mathbf{X}^*)\vartheta(\bar{\mathbf{S}}_i^*) \in [0,1], \tag{15}$$

where $\vartheta(\bar{\mathbf{S}}_i^*)$ denotes the minimum eigenvalue of $\bar{\mathbf{S}}_i^*$, obtained by solving (4)*. If $\mathbb{C}_i = \rho_i = 1$, then the solution of* (4) *is the optimum of the original non-relaxed problem* (13)*.*

The proposition gives a procedure where each node can check, locally and in real-time, the performance of its optimisation process and assess it by finding the optimal value for Problem 2. This is of key importance not only to verify the optimality of the estimation, but also the certificate can be used in other parts of the control pipeline to, e.g., assure optimality of the controller. In our case, we will leverage this process to design an ET communication rule.

V. STABILITY AND OPTIMALITY

In this section we analyse the stability and optimality of ECO-DKF. For clarity, we assume that all nodes broadcast their information to neighbours at all instants, such that the underlying communication graph is undirected and connected. In the next section we will relax the assumption.

Define $\eta_i = \hat{\mathbf{x}}_i - \mathbf{x}$,

$$\eta = [\eta_1^T, \ldots, \eta_N^T]^T, \quad \mathbf{H} = \text{block-diag}(\mathbf{H}_1, \ldots, \mathbf{H}_N)$$
$$\mathbf{z} = [\mathbf{z}_1^T, \ldots, \mathbf{z}_N^T]^T, \quad \mathbf{R} = \text{block-diag}(\mathbf{R}_1, \ldots, \mathbf{R}_N),$$
$$\mathbf{y} = [\mathbf{y}_1^T, \ldots, \mathbf{y}_N^T]^T, \quad \mathbf{U} = \text{block-diag}(\mathbf{U}_1, \ldots, \mathbf{U}_N),$$
$$\mathcal{A} = \mathbf{I} \otimes \mathbf{A}, \quad \hat{\mathbf{P}}^* = \text{block-diag}(\hat{\mathbf{P}}_1^*, \ldots, \hat{\mathbf{P}}_N^*),$$
$$\mathcal{Q} = \mathbf{I} \otimes \mathbf{Q}, \quad \bar{\mathbf{P}}^* = \text{block-diag}(\bar{\mathbf{P}}_1^*, \ldots, \bar{\mathbf{P}}_N^*),$$

and

$$\mathbf{Y} \overset{Eq.~(9)}{=} \mathbf{N}_w \mathbf{U} \text{ , where } \begin{array}{l} \lfloor \mathbf{N}_w \rfloor_{ij} = \mathbf{I} \text{ if } j \in \mathcal{J}_i \\ \lfloor \mathbf{N}_w \rfloor_{ij} = 0 * \mathbf{I} \text{ otherwise} \end{array} . \tag{16}$$

To demonstrate the global asymptotic stability of ECO-DKF, we adapt and prove two Lemmas from [4], extending them from the discrete-time centralized Kalman filter setting to our ECO-DKF setting. The first builds some useful matrices.

Lemma 1 (**Adapted from Lemma 2 in [4]**). *Given Eqs.* (10)-(11)*, the following holds:*

1) $\mathbf{F} = \mathbf{I} - \hat{\mathbf{P}}^*\mathbf{Y} = \hat{\mathbf{P}}^*(\bar{\mathbf{P}}^*)^{-1}$*.*

2) $\hat{\mathbf{P}}^* = \mathbf{FGF}^T$ *with* $\mathbf{G} = \mathcal{A}\hat{\mathbf{P}}^*\mathcal{A}^T + \mathcal{Q} + \mathbf{TR}^{-1}\mathbf{T}^T$ *and* $\mathbf{T} = \bar{\mathbf{P}}^*\mathbf{N}_w\mathbf{H}^T$*.*

Proof. Let $\mathbf{F} = \mathbf{I} - \hat{\mathbf{P}}^*\mathbf{Y}$, by Eq. (10) we can write $\hat{\mathbf{P}}^* = ((\bar{\mathbf{P}}^*)^{-1} + \mathbf{Y})^{-1}$. Then,

$$\hat{\mathbf{P}}^*((\bar{\mathbf{P}}^*)^{-1} + \mathbf{Y}) = \mathbf{I} \Rightarrow \mathbf{I} - \hat{\mathbf{P}}^*\mathbf{Y} = \hat{\mathbf{P}}^*(\bar{\mathbf{P}}^*)^{-1} \tag{17}$$

and the first statement is proved. Regarding the second statement, notice that Eq. (11) can be stacked as follows

$$\hat{\mathbf{x}}^* = \bar{\mathbf{x}}^* + \hat{\mathbf{P}}^*(\mathbf{y} - \mathbf{Y}\bar{\mathbf{x}}^*). \tag{18}$$

Then, taking into account that $\mathbf{y} = \mathbf{N}_w\mathbf{H}^T\mathbf{R}^{-1}\mathbf{z}$ and $\mathbf{Y} = \mathbf{N}_w\mathbf{U} = \mathbf{N}_w\mathbf{H}^T\mathbf{R}^{-1}\mathbf{H}$, Eq. (18) is rewritten from its information form as

$$\hat{\mathbf{x}}^* = \bar{\mathbf{x}}^* + \hat{\mathbf{P}}^*\mathbf{N}_w\mathbf{H}^T\mathbf{R}^{-1}(\mathbf{z} - \mathbf{H}\bar{\mathbf{x}}^*) = \bar{\mathbf{x}}^* + \mathbf{K}(\mathbf{z} - \mathbf{H}\bar{\mathbf{x}}^*) \tag{19}$$

with $\mathbf{K} = \hat{\mathbf{P}}^*\mathbf{N}_w\mathbf{H}^T\mathbf{R}^{-1}$ the Kalman Gain of the standard KF formulation. Let now consider the update of $\hat{\mathbf{P}}^*$, where by update, denoted as $(\hat{\mathbf{P}}^*)^+$, we refer to the operations that compute $\hat{\mathbf{P}}^*(k)$ from $\hat{\mathbf{P}}^*(k-1)$ and $\bar{\mathbf{P}}^*(k-1)$

$$(\hat{\mathbf{P}}^*)^+ = \mathbf{F}\bar{\mathbf{P}}^+\mathbf{F}^T + \mathbf{KRK}^T, \tag{20}$$

which comes from the standard KF [4]. By substituting $\mathbf{K}$ we obtain

$$(\hat{\mathbf{P}}^*)^+ = \mathbf{F}(\mathcal{A}\hat{\mathbf{P}}^*\mathcal{A}^T + \mathcal{Q})\mathbf{F}^T + \hat{\mathbf{P}}^*\mathbf{N}_w\mathbf{H}^T\mathbf{R}^{-1}\mathbf{H}\mathbf{N}_w^T(\hat{\mathbf{P}}^*)^T. \tag{21}$$

Given statement 1 of the Lemma,

$$(\hat{\mathbf{P}}^*)^+ = \mathbf{F}(\mathcal{A}\hat{\mathbf{P}}^*\mathcal{A}^T + \mathcal{Q} + \bar{\mathbf{P}}^*\mathbf{N}_w\mathbf{H}^T\mathbf{R}^{-1}\mathbf{H}\mathbf{N}_w^T(\bar{\mathbf{P}}^*)^T)\mathbf{F}^T, \tag{22}$$

and statement 2 is proved. $\qquad\square$

The next Lemma proves global asymptotic stability for general dynamics, used later to prove ECO-DKF global asymptotic stability.

Lemma 2 (**Adapted from Lemma 3 in [4]**). *Suppose that the error dynamics without noise are $\eta^+ = \mathbf{F}\mathcal{A}\eta$, with $\mathbf{F}$ as in Lemma 1. Then, the error dynamics is globally asymptotically stable system with a Lyapunov function $V(\eta) = \eta^T(\hat{\mathbf{P}}^*)^{-1}\eta$, provided that $g_i(\boldsymbol{\lambda}_i^*) = 1$ for all i, k and $G(k)$ is undirected and connected for all k.*

Proof. Given $V(\eta) = \eta^T(\hat{\mathbf{P}}^*)^{-1}\eta$ as Lyapunov function candidate,

$$\begin{aligned} \delta V =& (\eta^+)^T((\hat{\mathbf{P}}^*)^+)^{-1}\eta^+ - \eta^T(\hat{\mathbf{P}}^*)^{-1}\eta = \\ & \eta^T \left(\mathcal{A}^T\mathbf{F}^T((\hat{\mathbf{P}}^*)^+)^{-1}\mathbf{F}\mathcal{A} - (\hat{\mathbf{P}}^*)^{-1}\right)\eta = \\ & \eta^T \left(\mathcal{A}^T\mathbf{G}^{-1}\mathcal{A} - (\hat{\mathbf{P}}^*)^{-1}\right)\eta = \\ & -\eta^T \left((\hat{\mathbf{P}}^*)^{-1} - \mathcal{A}^T(\mathcal{A}\hat{\mathbf{P}}^*\mathcal{A}^T + \mathbf{W})^{-1}\mathcal{A}\right)\eta = \\ & -\eta^T\Lambda\eta, \end{aligned} \tag{23}$$

with $\mathbf{W} = \mathcal{Q} + \mathbf{TR}^{-1}\mathbf{T}^T \succ \mathbf{0}$ and $\Lambda = (\hat{\mathbf{P}}^*)^{-1} - \mathcal{A}^T(\mathcal{A}\hat{\mathbf{P}}^*\mathcal{A}^T + \mathbf{W})^{-1}\mathcal{A}$. The rest of the proof directly follows from Lemma 3 of [4], showing that $\Lambda \succ \mathbf{0}$. $\qquad\square$

Finally, to prove stability of ECO-DKF we employ a Lemma, whose proof can be found in [33], which allows to rewrite ECO-DKF filter equations and show that the optimum of optimization problem (4) is such that inequality (4b) is an equality.

Lemma 3. *Given optimisation problem* (4)*, $\bar{\mathbf{S}}_i^* = \sum_{j \in \mathcal{J}_i} \lambda_{ij}^*\bar{\mathbf{S}}_j$, for all i.*

Lemma 3 says that the optimisation in (4) becomes a standard discrete-time consensus protocol tuned to optimise the trace of the final consensus value of each node. Interestingly, the optimisation will implicitly assign largest λ_{ij}^* to nodes equipped with better sensors and estimates, which is a positive side effect.

These ingredients lead to the stability of the filter.

Theorem 1. *Algorithm 1 is a globally asymptotically stable estimator if $g_i(\boldsymbol{\lambda}_i^*) = 1$ for all i, k.*

Proof. First, rewrite Eq. (11) like Eq. (12). The aggregated prediction $\bar{\mathbf{x}}_i^*$ in Eq. (11) is

$$\bar{\mathbf{x}}_i^* = \bar{\mathbf{P}}_i^* \sum_{j \in \mathcal{J}_i} \lambda_{ij}^* \bar{\mathbf{P}}_j^{-1} \bar{\mathbf{x}}_j. \tag{24}$$

This expression can be rewritten as

$$\bar{\mathbf{x}}_i^* = \bar{\mathbf{P}}_i^* \sum_{j \in \mathcal{J}_i} \lambda_{ij}^* \bar{\mathbf{P}}_j^{-1} \bar{\mathbf{x}}_i + \bar{\mathbf{P}}_i^* \sum_{j \in \mathcal{N}_i} \lambda_{ij}^* \bar{\mathbf{P}}_j^{-1} (\bar{\mathbf{x}}_j - \bar{\mathbf{x}}_i). \tag{25}$$

From Lemma 3, $\bar{\mathbf{P}}_i^* \sum_{j \in \mathcal{J}_i} \lambda_{ij}^* \bar{\mathbf{P}}_j^{-1} = \mathbf{I}$. Using (25) in (11) gives

$$\hat{\mathbf{x}}_i^* = \bar{\mathbf{x}}_i + \hat{\mathbf{P}}_i^*(\mathbf{y}_i - \mathbf{Y}_i\bar{\mathbf{x}}_i) + (\mathbf{I} - \hat{\mathbf{P}}_i^*\mathbf{Y}_i)\bar{\mathbf{P}}_i^* \sum_{j \in \mathcal{N}_i} \lambda_{ij}^* \bar{\mathbf{P}}_j^{-1} (\bar{\mathbf{x}}_j - \bar{\mathbf{x}}_i). \tag{26}$$

Eq. (26) is equivalent to Eq. (12) but weighting each term with the result of optimising (4). Then, the noiseless dynamics of η_i is

$$\eta_i^+ = \mathbf{A}\eta_i - \hat{\mathbf{P}}_i^*\mathbf{Y}_i\mathbf{A}\eta_i + (\mathbf{I} - \hat{\mathbf{P}}_i^*\mathbf{Y}_i)\bar{\mathbf{P}}_i^* \sum_{j \in \mathcal{N}_i} \lambda_{ij}^* \bar{\mathbf{P}}_j^{-1}\mathbf{A}(\eta_j - \eta_i) \tag{27}$$

which can also be written in compact form for the whole network as

$$\eta^+ = \mathbf{F}\,\mathbf{L}_w\,\mathcal{A}\eta. \tag{28}$$

where $(\cdot)^+$ is the update operator defined as in Lemma 2 and the proof of Lemma 1, $\mathbf{F}$ is defined as in Lemma 1, and $\mathbf{L}_w$ is such that $\lfloor \mathbf{L}_w \rfloor_{ij} = \bar{\mathbf{P}}_i^* \lambda_{ij}^* \bar{\mathbf{P}}_j^{-1}$ for all $j \in \mathcal{J}_i$ and $\mathbf{0}$ otherwise. The last expression is similar to the error dynamics proved as globally asymptotically stable in Lemma 2, but with $\mathbf{L}_w$ in between. In the proof of Lemma 2 it is shown that

$$\Lambda = (\hat{\mathbf{P}}^*)^{-1} - \mathcal{A}^T\mathbf{G}^{-1}\mathcal{A} \succ \mathbf{0}. \tag{29}$$

Instead, we have that $\Lambda' = (\hat{\mathbf{P}}^*)^{-1} - \mathcal{A}^T\mathbf{L}_w^T\mathbf{G}^{-1}\mathbf{L}_w\mathcal{A}$. If

$$\mathcal{A}^T\mathbf{G}^{-1}\mathcal{A} \succeq \mathcal{A}^T\mathbf{L}_w^T\mathbf{G}^{-1}\mathbf{L}_w\mathcal{A} \quad \Rightarrow \quad \mathbf{I} \succeq \mathbf{L}_w, \tag{30}$$

then $\Lambda' \succ \mathbf{0}$ and global asymptotic stability is proved. Given Lemma 3, $\mathbf{L}_w$ is a row-stochastic matrix. Therefore, by linear algebra results, the absolute value of any eigenvalue of $\mathbf{L}_w$ is less than or equal to 1. This means that the eigenvalues of matrix $\mathbf{L}_w - \mathbf{I}$ are all negative or equal to 0 and

$$\mathbf{0} \succeq \mathbf{L}_w - \mathbf{I} \Rightarrow \mathbf{I} \succeq \mathbf{L}_w. \tag{31}$$

Thus, (30) holds and $\Lambda' \succ \mathbf{0}$, concluding the proof. $\square$

Notice that stability holds independently on certification. The next step is to demonstrate optimality under positive certification.

Theorem 2. *Assume that the solution of* (4) *is certified as optimal. Then, Algorithm 1 is optimal in the sense of Problem 1 and* (4) *provides the optimal consensus gain of Eq. (12).*

Proof. We reformulate Problem 1 as an optimal Bayesian estimation problem over the network $\mathcal{G}$. Let $Z(k) = \{\mathbf{z}(0), \ldots, \mathbf{z}(k)\}$, where $\mathbf{z} = [\mathbf{z}_1^T, \ldots, \mathbf{z}_N^T]^T$. The posterior that minimises the MSE is given by $\hat{\mathbf{x}}(k) = \mathrm{E}[\mathbf{x}(k)|Z(k)] = \int \mathbf{x}(k) P(\mathbf{x}(k)|Z(k))d\mathbf{x}(k)$. To obtain $P(\mathbf{x}(k)|Z(k))$ from $P(\mathbf{x}(k-1)|Z(k-1))$, the Bayesian prediction and correction equations are:

$$P(\mathbf{x}(k)|Z(k-1)) = \int P(\mathbf{x}(k)|\mathbf{x}(k-1)) P(\mathbf{x}(k-1)|Z(k-1))d\mathbf{x}(k-1)$$

$$P(\mathbf{x}(k)|Z(k)) = \frac{P(\mathbf{z}(k)|\mathbf{x}(k)) P(\mathbf{x}(k)|Z(k-1))}{\int P(\mathbf{z}(k)|\mathbf{x}(k)) P(\mathbf{x}(k)|Z(k-1))d\mathbf{x}(k)} \tag{32}$$

Under unknown correlations, $P(\mathbf{x}(k)|Z(k-1)) = \mathcal{N}(\bar{\mathbf{x}}(k), \bar{\mathbf{P}}(k))$:

$$\bar{\mathbf{x}}(k) = \mathcal{A}\hat{\mathbf{x}}^*(k-1) \quad \text{and} \quad \bar{\mathbf{P}}(k) = \mathcal{A}\hat{\mathbf{P}}^*(k-1)\mathcal{A}^T + \mathcal{Q}, \tag{33}$$

with $\hat{\mathbf{P}}^*(k-1) = \mathrm{block\text{-}diag}(\hat{\mathbf{P}}_1^*(k-1), \ldots, \hat{\mathbf{P}}_N^*(k-1))$. This is the prediction step. Regarding the correction step, first, we have that $\mathbf{y} = [\mathbf{y}_1^T, \ldots, \mathbf{y}_N^T]^T$, $\mathcal{R} = \mathrm{block\text{-}diag}(\mathbf{Y}_1^{-1}, \ldots, \mathbf{Y}_N^{-1})$, and $P(\mathbf{z}(k)|\mathbf{x}(k)) = \mathcal{N}(\mathcal{H}\mathbf{x}(k), \mathcal{R})$, where $\mathcal{H}$ is the block diagonal matrix such that $\mathbf{Y}^{-1}\mathbf{y}(k) = \mathcal{H}\mathbf{x}(k)$. On the other hand, $P(\mathbf{x}(k)|Z(k)) = \mathcal{N}(\bar{\mathbf{x}}^*(k), \bar{\mathbf{P}}^*(k))$, where $\bar{\mathbf{x}}^*(k)$ and $\bar{\mathbf{P}}^*(k)$ are the fused predictions at each node. In particular, given the restrictions in Problem 1, the fused prediction that ensures consistency (constraint 3) under unknown correlations (constraint 1) and one-hop communications (constraint 2) is given by the outer Löwner-John (LJ) method [36]. Among the different measures of ellipsoid size, $\mathrm{Tr}(\bar{\mathbf{S}}^{-1})$ in problem (4) minimises the MSE because MSE $= \mathrm{E}\left[\frac{1}{N}\sum_{i=1}^N ||\hat{\mathbf{x}}_i^* - \mathbf{x}||^2\right] = \mathrm{Tr}(\mathrm{block\text{-}diag}(\hat{\mathbf{P}}_1^*, \ldots, \hat{\mathbf{P}}_N^*)) = \mathrm{Tr}(\hat{\mathbf{P}}^*) = \mathrm{Tr}((\hat{\mathbf{S}}^*)^{-1})$. By assumption of the theorem, the solution of (4) is optimal, so the associated unbiased optimal fused prediction at each node is $\bar{\mathbf{x}}_i^* = \bar{\mathbf{P}}_i^* \sum_{j \in \mathcal{J}_i} \lambda_{ij}^* \bar{\mathbf{P}}_j^{-1} \bar{\mathbf{x}}_j$, which is equal to Eq. (5).

Using the fundamental Gaussian identity (Appendix D in [39]), $\mathcal{N}(\mathcal{H}\mathbf{x}(k), \mathcal{R})\mathcal{N}(\bar{\mathbf{x}}^*(k), \bar{\mathbf{P}}^*(k)) = \mathcal{N}(\mathcal{H}\bar{\mathbf{x}}^*(k), \mathcal{R} + \mathcal{H}\bar{\mathbf{P}}^*(k)\mathcal{H}^T)\mathcal{N}(\bar{\mathbf{x}}^*(k), ((\bar{\mathbf{P}}^*(k))^{-1} + \mathcal{H}\mathcal{R}^{-1}\mathcal{H}^T)^{-1})$. Besides, the integral $\int \mathcal{N}(\mathcal{H}\mathbf{x}(k), \mathcal{R})\mathcal{N}(\bar{\mathbf{x}}^*(k), \bar{\mathbf{P}}^*(k))d\mathbf{x}(k) = \mathcal{N}(\mathcal{H}\bar{\mathbf{x}}^*(k), \mathcal{R} + \mathcal{H}\bar{\mathbf{P}}^*(k)\mathcal{H}^T)$, and, therefore, $P(\mathbf{x}(k)|Z(k)) = \mathcal{N}(\hat{\mathbf{x}}^*(k), \hat{\mathbf{P}}^*(k))$. Finally, by definition, $(\hat{\mathbf{P}}^*(k))^{-1}\hat{\mathbf{x}}^*(k) = (\bar{\mathbf{P}}^*(k))^{-1}\bar{\mathbf{x}}^*(k) + \mathcal{H}\mathcal{R}^{-1}\mathbf{z}(k)$, and $\hat{\mathbf{P}}^*(k)(\hat{\mathbf{P}}^*(k))^{-1}\hat{\mathbf{x}}^*(k) = \hat{\mathbf{x}}^*(k) = \bar{\mathbf{x}}^*(k) + \mathbf{K}(k)(\mathbf{z}(k) - \mathcal{H}\bar{\mathbf{x}}^*(k))$. Using Lemma 1, the update/correction operations are written in information form:

$$\begin{aligned} \hat{\mathbf{P}}^*(k) &= ((\bar{\mathbf{P}}^*(k))^{-1} + \mathcal{R}^{-1})^{-1} \\ \hat{\mathbf{x}}^*(k) &= \bar{\mathbf{x}}^*(k) + \bar{\mathbf{P}}^*(k)(\mathbf{y}(k) - \mathcal{R}^{-1})\bar{\mathbf{x}}^*(k)). \end{aligned} \tag{34}$$

Equations (33) and (34), together with optimisation problem (4), are the ECO-DKF filter in Algorithm 1. Besides, the expression in Eq. (26) can be reformulated as

$$\hat{\mathbf{x}}_i^* = \bar{\mathbf{x}}_i + \hat{\mathbf{P}}_i^*(\mathbf{y}_i - \mathbf{Y}_i\bar{\mathbf{x}}_i) + \gamma\hat{\mathbf{P}}_i^* \sum_{j \in \mathcal{N}_i} \lambda_{ij}^* \bar{\mathbf{P}}_j^{-1}(\bar{\mathbf{x}}_j - \bar{\mathbf{x}}_i) \tag{35}$$

with $\gamma = (\mathbf{I} - \hat{\mathbf{P}}_i^*\mathbf{Y}_i)\bar{\mathbf{P}}_i^*(\hat{\mathbf{P}}_i^*)^{-1}$. The latter is equivalent to Eq. (12) but weighting each term with the result of optimising (4).

Therefore, Algorithm 1 is optimal in the sense of Problem 1 and (4) provides the optimal consensus gain of Eq. (12). $\square$

VI. EVENT-TRIGGERED SCHEME

In this section we describe the proposed ET general scheme and we study its theoretical properties to show that the stability and optimality results are preserved under PJC of the network.

A. Description of the ET scheme

Under an ET scheme, a node should communicate when it possesses relevant information for the network. Our ET rule leverages the output of (4) to decide the triggering. Remember that (4) finds the optimal combination of estimates under positive certification. In particular, $\boldsymbol{\lambda}_i^*$ weights how important is each prediction in the optimal fused estimate. If at instant k, $\max(\boldsymbol{\lambda}_i^*) = \lambda_{ii}^*$, then the greatest contribution to $\bar{\mathbf{x}}_i^*$ is $\bar{\mathbf{x}}_i$, so the information at node i is the best in its neighbourhood and should be broadcast. The result is the ET rule in Eq. (6).

Some aspects must be underlined. On the one hand, at instant k, node i receives $|\mathcal{N}_i(k)|$ messages from broadcasting neighbours. Thus, the network topology is time-varying. On the other hand,

two situations are possible in the ET rule at node i and instant k: $|\mathcal{N}_i(k)| > 0$ and $|\mathcal{N}_i(k)| = 0$. The former is the core of the ET rule because it determines the communication usage. The latter, despite influencing in the bandwidth as well, must be designed to avoid that the network converges to N individual KFs. As we will see later, this is key to assure stability.

B. Theoretical analysis

The two main properties analysed in Section V are stability and optimality. Regarding stability, the next Theorem shows that ECO-DKF is globally asymptotically stable under PJC.

Theorem 3. *ECO-DKF is global asymptotically stable for any time-varying graph sequence that satisfies PJC with $t < \infty$.*

Proof. The error dynamics η^+ in Theorem 1 now depends on time

$$\eta^+ = \mathbf{F}(k)\ \mathbf{L}_w(k)\ \mathcal{A}\eta, \tag{36}$$

From Theorem 1, since $\mathbf{L}_w(k)$ is row stochastic for any graph, the absolute value of its eigenvalues is always less than or equal to 1. Besides, $\mathbf{F}(k)$ is always a positive definite matrix and $\mathbf{I} \succeq \mathbf{F}(k)$ according to its definition in Lemma 1. Therefore, as a worst case scenario, we can consider $\mathbf{L}_w(k) = \mathbf{I} \Rightarrow \eta^+ = \mathbf{F}(k)\mathcal{A}\eta$.

Let j be one of the nodes. The error dynamics, η_j^+, are then determined by the eigenvalues and eigenvectors of $\mathbf{F}_j(k)\mathbf{A}$. The convergence rate of the KF is upper bounded by its steady-state gain $\mathbf{K}^{ss}$ [40], so let instead consider

$$\eta_j^+ = \mathbf{F}_j^{ss}\mathbf{A}\eta_j. \tag{37}$$

Let $\boldsymbol{v}_j = \{v_j^1, \ldots, v_j^r, \ldots, v_j^n\}$ and $\mathcal{V}_j = \{\mathbf{v}_j^1, \ldots, \mathbf{v}_j^r, \ldots, \mathbf{v}_j^n\}$ be the eigenvalues and eigenvectors of $\mathbf{F}_j^{ss}\mathbf{A}$. The initial conditions are expressed as $\eta_j(0) = \sum_{r=1}^{n} \gamma_j^r \mathbf{v}_j^r$ and the norm of the error dynamics in Eq. (37) can be upper bounded by $||\eta_j^+(k)|| \leq \sum_{r=1}^{n} || \left(v_j^r\right)^k \gamma_j^r \mathbf{v}_j^r||$. Network observability implies that for each node j and eigenvalue v_j^r, there exist at least a node j' such that $\kappa|\Re(v_j^r)| < 1$, with $\kappa \in \mathbb{R}_{>0}$ and $\Re(v_j^r)$ the real part of v_j^r. Otherwise, it would not be possible to reconstruct the state, even if the network was fully connected. In the worst case, due to the PJC with period equal to t, the estimation of node j' associated to v_j^r is updated through $t(N-1)$ ECO-DKF instants before reaching node j. At each instant, ECO-DKF corrects the estimates in two main phases: the optimisation in (4) and the KF steps. The former consists in finding the intersection of $|\mathcal{N}_i(k) \cup \{i\}|$ ellipsoids. Since the intersection can not be larger than the smallest ellipsoid, then the result is upper bounded by the best estimate among estimates at instant k. Thus, in the worst case the estimation error associated to v_j^r remains that of j'. On the other hand, in the worst case, the estimation error associated to v_j^r is only predicted in the KF steps without using the measurements to perform the correction step of the filter, and the estimation error associated to v_j^r increases as a function of the eigenvalues of $\mathbf{A}$. Therefore, when the estimation of node j' associated to v_j^r reaches node j, the error associated to v_j^r is, in the worst case, the error for j' increased by a *constant* value that depends on the eigenvalues of $\mathbf{A}$ and $t(N-1)$.

Since the estimation error associated to v_j^r in j' tends to zero when $k \to \infty$, then so does the error in node j. Applying this result to all the eigenvectors and nodes in the network, the result is proved. $\square$

The next Proposition addresses optimality.

Proposition 2. *Theorem 2 holds under PJC and network-observability conditions.*

Proof. At instant k, node i receives $|\mathcal{N}_i(k)|$ messages and conducts optimisation (4) and certification (14) steps. The case where $|\mathcal{N}_i(k)| = 0$ is trivial and $\bar{\mathbf{x}}_i^*(k) = \bar{\mathbf{x}}_i(k)$ and $\bar{\mathbf{P}}_i^*(k) = \bar{\mathbf{P}}_i(k)$, according to restrictions 1 and 2 in Problem 1. In the general case, certifiability and optimality in Proposition 1 and Theorem 2 are proved for any $|\mathcal{N}_i(k)|$. Therefore, Theorem 2 holds irrespective of the connectivity and observability of the network. $\square$

The previous results are interesting not only because they demonstrate that ECO-DKF is still optimal and stable under PJC, but also because this can be done without requiring any additional information. Regarding the ET rule (6), by letting each node to communicate with probability $p > 0$ when it does not receive information, we force that nodes do not isolate forever, and ET rule (6) guarantees PJC with an expected value of t bounded.

To complete the solution, we show that ECO-DKF is a consistent estimator under PJC conditions.

Proposition 3. *ECO-DKF is a consistent estimator under PJC conditions.*

Proof. Recalling the definition of (4), the outer LJ method gives a fused $\bar{\mathbf{S}}_i^*$ which is always consistent. Moreover, considering that we assume unbiased sensors, $\bar{\mathbf{x}}_i^*$ preserves consistency. This is because: (i) constraint (4c) ensures positive weights and their sum less or equal to 1, so it is a non-increasing combination of estimates; (ii) Lemma 3 shows that the sum of the weights is always equal to 1, so it is a non-decreasing combination of estimates. Thus, the aggregation of estimates in (5) is also unbiased. Besides, Theorem 3 states that ECO-DKF is globally asymptotically stable under PJC conditions, which means that once estimates are aggregated, the update and prediction steps preserve consistency as well. Therefore, the whole algorithm is consistent. $\square$

VII. ILLUSTRATIVE RESULTS

We first evaluate the performance of the Time-Triggered (TT) ECO-DKF, i.e., ECO-DKF with $g_i(\boldsymbol{\lambda}_i^*(k-1), k) = 1$ for all i, k, to verify the stability and optimality properties. Then, different ET rules are compared, studying their impact in the communication bandwidth and convergence speed.

A. Time-triggered ECO-DKF vs Time-triggered DKFs

We compare ECO-DKF in the TT setting with respect to other TT DKFs: (i) *AtA-ECO-DKF*, our proposal but assuming that $(i, j) \in E(k) \ \forall i \neq j, k$, (ii) *OCDFK*, Algorithm 3 from [3] with consensus gain $\gamma_i(k) = 10^{-4}/(1 + ||\mathbf{M}_i(k)||_F)$ to ensure stability, with $|| \cdot ||_F$ the Frobenius norm, (iii) *TCDFK*, Algorithm 2 from [41] with one consensus step for a fair comparison, (iv) *HDfKF*, complete algorithm in [12], (v) *HADfKF*, simplified algorithm in [12], (vi) *CKF*, centralised equivalent KF. The target system is the same 2D particle in a circular orbit used in [33], with the same parameters.

We analyse two scenarios. Experiment 1 initialises a random sensor network, with appropriate parameters to obtain a sparse connected topology. Then, a random uniform distribution decides the quantities sensors measure among two options: measuring x or y. Note that local observability never holds, to test ECO-DKF in a network-observability setting. $\mathbf{H}_i$ is picked uniformly in the range $[1, 3]$. A Bernoulli distribution with $p = 0.5$ decides the diagonal of $\mathbf{R}_i$ in the range $[3, 5] \times 10^{-2}$ or $[3, 5]$, i.e., high-quality or low-quality. This is done 100 times, computing the averaged Mean Square Error MSE$:= \frac{1}{100} \sum_{i=1}^{N} \mathrm{E}[||\hat{\mathbf{x}}_i - \mathbf{x}||^2]$ over the experiments as in [14] and [12]. Experiment 2 is the same as Experiment 1, but only one sensor is of high quality, and sensors now can measure x, y or both.

The results of Experiments 1 are shown in Fig. 1a. The best performance among the distributed estimators is obtained by ECO-DKF, with a difference of more than an order of magnitude with the

other state-of-the-art filters in steady-state MSE. Besides, ECO-DKF is also the fastest filter, with the fastest asymptotic convergence to the CKF. The OCDKF achieves similar MSE performance than the AtA-ECO-DKF: adding more neighbours hinders the optimization problem in ECO-DKF. The other consensus-based DKF, TCDKF, is far from them because it needs multiple consensus steps within instants for a good performance [41]. The diffusion-based DKF, HDfKF and HADfKF, exhibit worse suboptimal performance as well. These differences also hold in Experiment 2, as it is shown in Fig. 1b. Therefore, ECO-DKF obtains the best performance among DKFs.

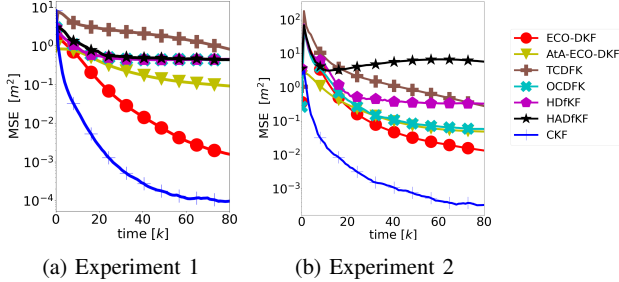(a) Experiment 1   (b) Experiment 2

Fig. 1: MSE for the different estimators in the TT case.

We have assessed the tightness of the relaxation by checking the certification values. A 99.33% of the times in Experiment 1 and 95.92% of the times in Experiment 2, $\mathbb{C}_i = \rho_i = 1$, so the SDP relaxation is *tight*. It is also noteworthy that the lower values of ρ are at 0.65, and this is only in the initialisation.

In relation with times, the computation of optimisations (4) and (14), run at each iteration, takes 50ms using a standard laptop. This is consistent with the election of the sample time $T = 100$ms. Problem (4) is solved in a polynomial number of iterations and arithmetic operations [42], so the computational cost scales with the number of neighbours. Whether it is preferable to compute more or communicate more will depend on the network resources, but in general, communicating multiple times may imply a total run time and risk of disturbances greater than ours. In this sense, we recall that ECO-DKF only communicates, at most, once per instant k.

B. Evaluation of event-triggered schemes in ECO-DKF

We now evaluate the impact of the ET rules in the NoB and MSE. We repeat Experiment 1 and 2 but with different ET rules: TT ECO-DKF (#O), the original TT ECO-DKF tested in Subsection VII-B; connected ECO-DKF (#C), which corresponds to ET rule (6) with $p(k) = 1 \ \forall k$; disconnected ECO-DKF (#D), which corresponds to ET rule (6) with $p(k) = 0 \ \forall k$; stochastic ECO-DKF (#S), which corresponds to ET rule (6) with $p(k) \sim \mathcal{NBB}(\alpha, \beta, n, k_l, k)$. Here, $\mathcal{NBB}(\alpha, \beta, n, k_l, k)$ is a normalised beta-binomial distribution whose shape is given by α, β, n. Given the shape, $p(k)$ increases when k_l increases, where $k_l \in \mathbb{N}$ is the number of instants that have passed since the last time $g_i(\boldsymbol{\lambda}_i^*) = 1$. The normalisation of the beta binomial is to force that, at some point, $p(k) = 1$, i.e., to force PJC. Finally, we also test a Jensen-Shannon (JS) divergence ECO-DKF (#J), which corresponds to a SoD rule based on the JS divergence

$$g_i(\boldsymbol{\lambda}_i^*) = \begin{cases} 1, & \text{if } D_{JS}(\bar{\mathbf{S}}_i^*(k) || \bar{\mathbf{S}}_i^*(k - k')) > \tau \\ 0, & \text{otherwise} \end{cases} \quad \forall i \in V \quad (38)$$

where τ is a tuned threshold, k' is the last time $g_i(\boldsymbol{\lambda}_i^*) = 1$ and JS is a symmetric version of the KL divergence. Note that ET rule (38) implicitly depends on $\boldsymbol{\lambda}_i^*$ since it is the divergence between the latest computation of (4) and the latest time node ET rule (38) was equal to 1. For #S, we have calibrated two combinations of parameters: $\alpha =$

$\{5, 2\}$, $\beta = \{1, 2\}$ and $n = \{10, 15\}$. For #J we have empirically tuned the threshold to be $\tau = 100$.

Figure 2 shows the evolution of NoB with time for the different ET rules, taking the average over the simulation runs. We can see that for #O, NoB $= N$ always, since it is a TT version of ECO-DKF. In contrast, #C reduces to a half the NoB, which is a significant decrease with a very simple tuning-free ET rule. As expected, #D evolves to 0 as the network tends to be disconnected. Regarding the stochastic rules (#S), depending on the shape of the beta-binomial distribution the network converges to a different steady-state NoB. This also affects the NoB transient, in the sense that a sharper distribution at lower k' values imply less time to trigger, reason why in #S2 there is a very low NoB before $k = 30$. Finally, the Jensen-Shannon trigger protocol achieves a trade-off between TT ECO-DKF (in the first instants) and a connected ECO-DKF (for $k > 60$).
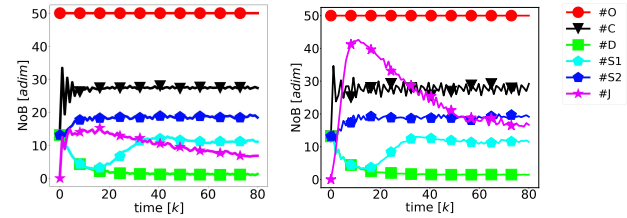

Fig. 2: Evolution of NoB average over simulation runs with time for the different ET rules: (left) Experiment 1, (right) Experiment 2.

Efficiency in communications is followed by a non-degraded performance. In Fig. 3 we depict the MSE of the ET rules in Experiments 1 and 2. In the first instants, those protocols with more communications have the worst MSE (e.g., #O) due to the poor initial estimates. However, afterwards, the tendency is inverted, (e.g., #J). In steady-state, the improvement in communications given by #S1 and #S2 is followed by a slight improvement in MSE. #C has a good trade-off between MSE and NoB. #D is the worst estimator; indeed, #D failed 10 runs in Experiment 1 and 13 runs in Experiment 2. From Fig. 1, the MSE of ECO-DKF is the closest to the CKF. In fact, given the high number of positive certifications, ECO-DKF is almost always optimal in the TT version. For general ET cases, ECO-DKF obtains almost the same performance compared to the TT version (Fig. 3), so the ET rule does not significantly affect the performance.
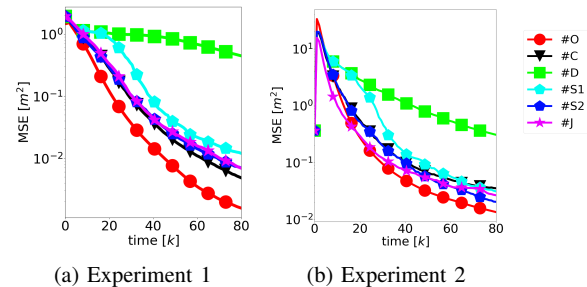

(a) Experiment 1   (b) Experiment 2

Fig. 3: MSE for the different estimators in the ET case.

	Experiment 1		Experiment 2
ET rule	Avg($\mathbb{C}_i^f = \rho_i^f = 1$)	ET rule	Avg($\mathbb{C}_i^f = \rho_i^f = 1$)
#O	99.33%	#O	95.92%
#C	79.30%	#C	77.74%
#D	11.80%	#D	12.11%
#S1	44.11%	#S1	44.14%
#S2	69.88%	#S2	69.16%
#J	47.69%	#J	67.24%

TABLE I: Certification results during the ET simulations.

Finally, in Table I we evaluate certifiability and optimality. To do so, those cases where $|\mathcal{N}_i(k)| = 0$ are removed from $\mathbb{C}_i$ and ρ_i since they are trivial to compute, denoted by $\mathbb{C}_i^f$ and ρ_i^f respectively. The results are averaged over the sensors of the network at each instant. Apart from #O, the only ET rule which resists with reasonable certification levels is #C: the disconnected rule hardly ever certifies as optimal; #S1 and #S2 have worse certification results; and #J has a large variability among experiments, so it is hard to predict its certification performance.

VIII. CONCLUSIONS

This paper has presented ECO-DKF, the first event-triggered and certifiable optimal DKF. It has solved two main problems regarding DKFs: optimality in the estimation and reduction of the communication bandwidth. The outer LJ fuses neighbouring estimates with certifiable guarantees of optimality. The output is integrated in an information DKF, achieving a stable filter on heterogeneous sensor networks, with minimal message size and no tuning. We have proved global asymptotic stability of the estimator and optimality under positive certification. Moreover, a novel ET theory has been derived from the outer LJ method output. Under PJC, ECO-DKF preserves the properties of the algorithm. The ET rule is inexpensive to compute, and avoids individual communication links and multiple information exchanges. ECO-DKF surpasses the state-of-the-art DKFs while decreasing the communication bandwidth usage.

REFERENCES

[1] R. E. Kalman, "A new approach to linear filtering and prediction problems," *ASME Journal of Basic Engineering*, vol. 82, no. 1, pp. 35–45, 1960.

[2] R. Olfati-Saber, "Distributed Kalman filter with embedded consensus filters," in *IEEE Conference on Decision and Control*, 2005, pp. 8179–8184.

[3] ——, "Distributed Kalman filtering for sensor networks," in *IEEE Conference on Decision and Control*, 2007, pp. 5492–5498.

[4] ——, "Kalman-consensus filter: Optimality, stability, and performance," in *IEEE Conference on Decision and Control*, 2009, pp. 7036–7042.

[5] S. Zhu, C. Chen, W. Li, B. Yang, and X. Guan, "Distributed optimal consensus filter for target tracking in heterogeneous sensor networks," *IEEE Transactions on Cybernetics*, vol. 43, no. 6, pp. 1963–1976, 2013.

[6] A. T. Kamal, J. A. Farrell, and A. K. Roy-Chowdhury, "Information weighted consensus filters and their application in distributed camera networks," *IEEE Transactions on Automatic Control*, vol. 58, no. 12, pp. 3112–3125, 2013.

[7] G. Battistelli, L. Chisci, G. Mugnai, A. Farina, and A. Graziano, "Consensus-based linear and nonlinear filtering," *IEEE Transactions on Automatic Control*, vol. 60, no. 5, pp. 1410–1415, 2014.

[8] S. Das and J. M. Moura, "Consensus+Innovations distributed Kalman filter with optimized gains," *IEEE Transactions on Signal Processing*, vol. 65, no. 2, pp. 467–481, 2016.

[9] S. Battilotti, F. Cacace, and M. d'Angelo, "A stability with optimality analysis of consensus-based distributed filters for discrete-time linear systems," *Automatica*, vol. 129, p. 109589, 2021.

[10] S. J. Julier and J. K. Uhlmann, "A non-divergent estimation algorithm in the presence of unknown correlations," in *IEEE American Control Conference*, vol. 4, 1997, pp. 2369–2373.

[11] S. Julier, "General decentralized data fusion with covariance intersection (CI)," *Handbook of multisensor data fusion*, 2001.

[12] J. Hu, L. Xie, and C. Zhang, "Diffusion Kalman filtering based on covariance intersection," *IEEE Transactions on Signal Processing*, vol. 60, no. 2, pp. 891–902, 2011.

[13] M. Reinhardt, B. Noack, P. O. Arambel, and U. D. Hanebeck, "Minimum covariance bounds for the fusion under unknown correlations," *IEEE Signal Processing Letters*, vol. 22, no. 9, pp. 1210–1214, 2015.

[14] F. S. Cattivelli and A. H. Sayed, "Diffusion strategies for distributed Kalman filtering and smoothing," *IEEE Transactions on Automatic Control*, vol. 55, no. 9, pp. 2069–2084, 2010.

[15] Z. Deng, P. Zhang, W. Qi, J. Liu, and Y. Gao, "Sequential covariance intersection fusion Kalman filter," *Information Sciences*, vol. 189, pp. 293–309, 2012.

[16] Y. Zhang, C. Wang, N. Li, and J. Chambers, "Diffusion Kalman filter based on local estimate exchanges," in *IEEE International Conference on Digital Signal Processing*, 2015, pp. 828–832.

[17] S. Boyd, L. El Ghaoui, E. Feron, and V. Balakrishnan, *Linear matrix inequalities in system and control theory*. SIAM, 1994.

[18] J. Wu, Q.-S. Jia, K. H. Johansson, and L. Shi, "Event-based sensor data scheduling: Trade-off between communication rate and estimation quality," *IEEE Transactions on Automatic Control*, vol. 58, no. 4, pp. 1041–1046, 2012.

[19] Y. S. Suh, V. H. Nguyen, and Y. S. Ro, "Modified Kalman filter for networked monitoring systems employing a send-on-delta method," *Automatica*, vol. 43, no. 2, pp. 332–338, 2007.

[20] W. Li, Y. Jia, and J. Du, "Event-triggered Kalman consensus filter over sensor networks," *IET Control Theory & Applications*, vol. 10, no. 1, pp. 103–110, 2015.

[21] Q. Liu, Z. Wang, X. He, and D.-H. Zhou, "Event-based distributed filtering with stochastic measurement fading," *IEEE Transactions on Industrial Informatics*, vol. 11, no. 6, pp. 1643–1652, 2015.

[22] C. Liu and S. Sun, "Event-triggered optimal and suboptimal distributed Kalman consensus filters for sensor networks," *Journal of the Franklin Institute*, vol. 358, no. 9, pp. 5163–5183, 2021.

[23] C. Zhang and Y. Jia, "Distributed Kalman consensus filter with event-triggered communication: Formulation and stability analysis," *Journal of the Franklin Institute*, vol. 354, no. 13, pp. 5486–5502, 2017.

[24] K. You and L. Xie, "Kalman filtering with scheduled measurements," *IEEE Transactions on Signal Processing*, vol. 61, no. 6, pp. 1520–1530, 2012.

[25] Q. Liu, Z. Wang, X. He, and D.-H. Zhou, "Event-based recursive distributed filtering over wireless sensor networks," *IEEE Transactions on Automatic Control*, vol. 60, no. 9, pp. 2470–2475, 2015.

[26] G. Battistelli, L. Chisci, and D. Selvi, "A distributed Kalman filter with event-triggered communication and guaranteed stability," *Automatica*, vol. 93, pp. 75–82, 2018.

[27] S. Trimpe and M. C. Campi, "On the choice of the event trigger in event-based estimation," in *IEEE International Conference on Event-based Control, Communication, and Signal Processing*, 2015, pp. 1–8.

[28] G. Battistelli, A. Benavoli, and L. Chisci, "Data-driven communication for state estimation with sensor networks," *Automatica*, vol. 48, no. 5, pp. 926–935, 2012.

[29] D. Han, Y. Mo, J. Wu, B. Sinopoli, and L. Shi, "Stochastic event-triggered sensor scheduling for remote state estimation," in *IEEE Conference on Decision and Control*, 2013, pp. 6079–6084.

[30] S. Trimpe and R. D'Andrea, "Event-based state estimation with variance-based triggering," *IEEE Transactions on Automatic Control*, vol. 59, no. 12, pp. 3266–3281, 2014.

[31] B. Sinopoli, L. Schenato, M. Franceschetti, K. Poolla, M. I. Jordan, and S. S. Sastry, "Kalman filtering with intermittent observations," *IEEE Transactions on Automatic Control*, vol. 49, no. 9, pp. 1453–1464, 2004.

[32] G. Battistelli, L. Chisci, L. Gao, and D. Selvi, "Event-triggered distributed Bayes filter," in *IEEE European Control Conference*, 2019, pp. 2731–2736.

[33] E. Sebastián, E. Montijano, and C. Sagüés, "All-in-one: Certifiable optimal distributed kalman filter under unknown correlations," in *IEEE Conference on Decision and Control*, 2021, pp. 6578–6583.

[34] H. Yang, J. Shi, and L. Carlone, "Teaser: Fast and certifiable point cloud registration," *IEEE Transactions on Robotics*, vol. 37, no. 2, pp. 314–333, 2020.

[35] D. Shi, T. Chen, and L. Shi, "On set-valued Kalman filtering and its application to event-based state estimation," *IEEE Transactions on Automatic Control*, vol. 60, no. 5, pp. 1275–1290, 2014.

[36] F. John, "Extremum problems with inequalities as subsidiary conditions," in *Traces and emergence of nonlinear programming*. Springer, 2014, pp. 197–215.

[37] B. D. Anderson and J. B. Moore, *Optimal filtering*. Courier Corporation, 2012.

[38] D. Henrion, S. Tarbouriech, and D. Arzelier, "LMI approximations for the radius of the intersection of ellipsoids: Survey," *Journal of Optimization Theory and Applications*, vol. 108, no. 1, pp. 1–28, 2001.

[39] R. P. Mahler, *Statistical multisource-multitarget information fusion*. Artech House Norwood, MA, USA, 2007, vol. 685.

[40] M. S. Grewal and A. P. Andrews, *Kalman filtering: Theory and Practice with MATLAB*. John Wiley & Sons, 2014.

[41] S. P. Talebi and S. Werner, "Distributed Kalman filtering and control through embedded average consensus information fusion," *IEEE Transactions on Automatic Control*, vol. 64, no. 10, pp. 4396–4403, 2019.

[42] Y. Nesterov and A. Nemirovskii, *Interior-point polynomial algorithms in convex programming*. SIAM, 1994.