Distributed Energy-Based Estimation Over Harvesting-Constrained Sensor Networks

Chen, Shuqi; Ho, Daniel W. C.

Published in:
IEEE Transactions on Cybernetics

Online published: 09/05/2023

Document Version:
Post-print, also known as Accepted Author Manuscript, Peer-reviewed or Author Final version

Publication record in CityU Scholars:
Go to record

Published version (DOI):
10.1109/TCYB.2023.3270872

Publication details:
https://doi.org/10.1109/TCYB.2023.3270872

Citing this paper
Please note that where the full-text provided on CityU Scholars is the Post-print version (also known as Accepted Author Manuscript, Peer-reviewed or Author Final version), it may differ from the Final Published version. When citing, ensure that you check and use the publisher's definitive version for pagination and other details.

General rights
Copyright for the publications made accessible via the CityU Scholars portal is retained by the author(s) and/or other copyright owners and it is a condition of accessing these publications that users recognise and abide by the legal requirements associated with these rights. Users may not further distribute the material or use it for any profit-making activity or commercial gain.

Publisher permission
Permission for previously published items are in accordance with publisher's copyright policies sourced from the SHERPA RoMEO database. Links to full text versions (either Published or Post-print) are only available if corresponding publishers allow open access.

Take down policy
Contact lbscholars@cityu.edu.hk if you believe that this document breaches copyright and provide us with details. We will remove access to the work immediately and investigate your claim.
Distributed Energy-Based Estimation Over Harvesting-Constrained Sensor Networks

Shuqi Chen, Daniel W. C. Ho, Fellow, IEEE

Abstract—This article investigates the distributed joint state and fault estimation issue for a class of nonlinear time-varying systems over sensor networks constrained by energy harvesting. It is assumed that data transmission between sensors requires energy consumption, and each sensor can harvest energy from the external environment. A Poisson process models the energy harvested by each sensor, and the sensor’s transmission decision depends on its current energy level. One can obtain the sensor transmission probability through a recursive calculation of the probability distribution of the energy level. Under such energy harvesting constraints, the proposed estimator only uses local and neighbor data to simultaneously estimate the system state and the fault, thereby establishing a distributed estimation framework. Moreover, the estimation error covariance is determined to possess an upper bound, which is minimized by devising energy-based filtering parameters. The convergence performance of the proposed estimator is analyzed. Finally, a practical example is presented to verify the usefulness of the main results.

Index Terms—Distributed energy-based estimation, energy harvesting sensors, sensor networks, state and fault estimation.

I. INTRODUCTION

Due to the rise of sensor network technology, considerable research efforts have been devoted to the distributed estimation/filtering problem [1]–[10]. One of the main features of distributed state estimation can be generalized as utilizing only the information of the sensor itself and its neighbors to acquire the local estimate.

It is well known that the energy issue has long been a crucial topic in sensor networks. As an energy replenishing scheme, energy harvesting technology replaces the original battery-powered sensors with sensors equipped with energy harvesters. The energy harvesters, including solar panels, windmills, and other devices, access energy from the external surroundings and convert it into electrical energy, which is stored in sensors for transmission. The energy harvesting technology has been applied in numerous academic and industrial fields (see [11]–[14]). For instance, ambient backscatter sensors that harvest energy from radio-frequency sources have been used in [13] to design decision fusion rules. In [14], the decentralized detection problem has been addressed in sensor networks composed of energy harvesting sensors. Furthermore, some encouraging related results have started to appear on the estimation problem. For example, the state estimation problem with an energy harvesting sensor has been studied for transmission energy scheduling/allocation [15]–[17], energy-dependent triggering protocol [18], and finite-horizon filter design [19].

Further investigations extending to the multisensor estimation problem can be found in [20], [21].

Regarding the traditional distributed estimation issue over sensor networks, the energy consumption for transmission between different sensors has not been fully considered. In practice, however, data transmission requires energy. The continuous communication will drain the battery power and cause the sensor to become idle due to insufficient energy. As a result, the system performance may be impaired. In this distributed framework, the energy harvesting technology provides a feasible approach to power sensors, thereby solving the energy consumption problem. Due to the randomness of the energy harvesting level, the energy amount stored in sensors also has random characteristics. The energy level held by sensors may need to be increased, thus failing to maintain communication between sensors at some time. Consequently, the adoption of energy harvesting technology will impose constraints on communication. To be specific, it will lead to intermittent data transmission. In the face of such discontinuous communication, the analysis of the transmission probability plays a vital role in the distributed estimator design. Very recently, distributed filtering schemes limited by energy harvesting have been developed in [22], [23] regarding linear multisensor systems. However, considering the intricate communication among sensors induced by energy harvesting constraints, the corresponding distributed estimation design for nonlinear systems still needs more attention.

In addition to the energy issue, faults are another unavoidable problem that may damage system performance. In actual cases, even small incipient faults gradually evolve into multitudinous abnormal conditions. Accordingly, fault diagnosis approaches are of great significance [24]. In general, it is demanding to extract accurate quantitative information on fault signals exclusively by applying fault detection and isolation methods [25], [26]. As such, the so-called fault estimation problem, which aims to obtain the magnitude or time-varying characteristics of fault signals, has attracted growing interest (see [27]–[32]). Nevertheless, as far as we know, the distributed joint state and fault estimation problem over sensor networks under energy harvesting constraints has not been dealt with, let alone that the system nonlinearity is also considered.

Inspired by the above discussion, this article addresses the distributed state and fault estimation issue concerning a class of nonlinear time-varying systems over energy-harvesting-
constrained sensor networks. The main contributions of this article are declared as follows.

1) Compared with the estimation/control schemes [15]–[21] under energy harvesting constraints, this article investigates intricate communication among multiple energy harvesting sensors in a distributed fashion.

2) Unlike the energy-related distributed estimation/control designs in [22], [23], and [33], a comprehensive system containing fault signals and time-varying nonlinearities is considered in this article. Hence, a new unified distributed estimator is proposed based on the energy-harvesting-constrained transmission protocol.

3) To handle the intermittent communication among sensors caused by the uncertainty of energy harvesting, a recursive calculation of the probability distribution of the energy amount is given to derive the transmission probability of each sensor.

4) Under the coupled effects of faults, system nonlinearity, and energy-constrained communication, an upper bound of the estimation error covariance is provided and minimized by designing energy-dependent filter parameters.

The article is organized as follows. The formulated problem is described in Section II. Under energy harvesting constraints, the computation of sensor transmission probability, the design of distributed state and fault estimator, and the estimator convergence analysis are studied in Section III. A numerical example is presented in Section IV to confirm the main results. Finally, conclusions are discussed in Section V.

**Notations:** $\mathbb{R}^n$ means the $n$-dimensional Euclidean space. Let $\mathbb{N}$ refer to the set of non-negative integers. $\| \cdot \|$ is the Euclidean norm of a vector or the spectral norm of a matrix. $X^T$ and $X^{-1}$ are the transpose and the inverse of matrix $X$, respectively. $\text{tr}\{\cdot\}$ denotes the trace. Let $\lambda_{\text{max}}(\cdot)$ be the largest eigenvalue. For real symmetric matrices $W$ and $V$, $W \succ V$ ($W \succeq V$) means $W - V$ is positive definite (positive semi-definite). Let $I_n$ and $0_{n \times m}$ be the $n \times n$ identity matrix and the $n \times m$ zero matrix, respectively. Sometimes $I$ and $0$ are used for simplicity. $\text{diag}\{\cdots\}$ is a block-diagonal matrix. $1_n$ is the $n$-dimension row vector whose elements are all 1s. $\text{Pr}(\cdot)$ means the probability. $\mathbb{E}\{\cdot\}$ is the expectation operator.

**II. PROBLEM FORMULATION**

A directed graph $G = (\mathcal{N}, \mathcal{E})$ depicts a sensor network, where $\mathcal{N} = \{1, 2, \ldots, n\}$ represents the set of sensors, and $\mathcal{E}$ means the set of edges. An edge $(i, j) \in \mathcal{E}$ implies that sensor $i$ can transmit information to sensor $j$. The self-loops are prohibited, that is, $(i, i) \notin \mathcal{E}$. The set of in-neighbors of sensor $i$ is expressed as $\mathcal{N}^i = \{j | (j, i) \in \mathcal{E}\}$. Let $\tilde{n}^i$ denote the total number of labels in $\mathcal{N}^i$. Consider the discrete nonlinear time-varying system as follows:

$$x_{k+1} = f_k(x_k) + B_k \gamma_k + w_k$$  

(1)

where $x_k \in \mathbb{R}^{n_x}$, $\gamma_k \in \mathbb{R}^{n_y}$, and $w_k \in \mathbb{R}^{n_w}$ are the system state, the fault signal, and the process noise, respectively, and $B_k$ is a known matrix with appropriate dimension. Assume that the initial state $x_0$ has the mean $\bar{x}_0$. The dynamic characteristics of the fault signal $\gamma_k$ are described by

$$\gamma_{k+1} = D_k \gamma_k$$  

(2)

where $D_k$ is a known time-varying matrix. The nonlinear function $f_k(\cdot)$ is supposed to satisfy

$$\|f_k(h) - f_k(g) - A_k(h - g)\| \leq \alpha_k \|h - g\|, \ \forall h, g \in \mathbb{R}^{n_x}$$  

(3)

with a known matrix $A_k$ and a non-negative scalar $\alpha_k$.

**Remark 1:** This work considers the distributed state and fault estimation problem for a nonlinear time-varying system (1) with additive fault. A general fault model (2) widely used in [29]–[32] is adopted. Assume $D_k$ in (2) is known since some priori information of the fault signal (e.g., incipient and constant faults) is available based on engineering background through existing fault detection methods.

**Remark 2:** The nonlinear description (3) with the time-varying matrix $A_k$ reveals the distance between the nonlinear system model (1) and its nominal linear model. Typically, the linearization methods evaluate the maximum deviation of the nonlinear model from the nominal linear model. In addition, (3) becomes the standard Lipschitz condition when $A_k = 0$. Thus, (3) is a general nonlinear description containing the Lipschitz condition as a particular case.

The $i$th sensor’s measurement model is

$$z_{k,i} = C_k^i x_k + v_{k,i}$$  

(4)

where $z_{k,i} \in \mathbb{R}^{n_z}$ is the measurement output of sensor $i$, and $v_{k,i} \in \mathbb{R}^{n_z}$ is the measurement noise. $C_k^i$ is a known matrix. $w_k$ in (1) and $v_{k,i}$ in (4) are zero-mean Gaussian white noise sequences with covariances $Q_k$ and $R_k^i$, respectively. The variables $x_0$, $w_k$, and $v_{k,i}$ are mutually uncorrelated.

Suppose that each data transmission between two adjacent sensors consumes $\Delta \in \mathbb{R}^{\text{unit}(s)}$ of energy. For the convenience of analysis, $\Delta$ is assigned to be the value one herein, and $N^i \triangleq n^i \Delta$, where $n^i$ is the number of sensor $i$’s out-neighbors. For ease of notation, the symbol $\Delta$ is dropped in the remaining paper. Let $E_k^i \in \{0, 1, \ldots, E^i\}$ indicate the energy level of sensor $i$ at time instant $k$, where $E^i$ is the maximum number of units of energy that can be stored by sensor $i$. As described in Fig. 1, the considered communication scenario is when energy is sufficient, that is, $E_k^i \geq N^i$, sensor $i$ performs data transmission to all its neighbors at time instant $k$ and consumes $N^i$ units of energy in total.

![Fig. 1: Sensor network subject to energy harvesting constraints.](image)
Mathematically speaking, an indicator variable $\beta_k^i$ representing the transmission decision is denoted by

$$\beta_k^i = \begin{cases} 1, & \text{if } E_k^i \geq N^i \\ 0, & \text{otherwise} \end{cases}$$

(5)

and the evolution of the energy level of sensor $i$ is

$$E_{k+1}^i = \min\{E_k^i - \beta_k^i N^i + H_k^i, E_k^i\}$$

(6)

where the initial energy level satisfies $E_0^i = E_k^i \leq \bar{E}^i$. Here, $H_k^i$ in (6) is the amount of energy harvested by sensor $i$ at time instant $k$, which is characterized as a Poisson process with mean $\lambda_k^i$ as follows:

$$\Pr(H_k^i = q) = \left(\frac{\lambda_k^i}{q!}\right)^q e^{-\lambda_k^i}, \quad q \in \mathbb{N}$$

(7)

where $\Pr(H_k^i = q)$ is abbreviated as $p_{k,q}^i$ herein, and $\lambda_k^i$ represents the energy harvesting rate, that is, the average number of units of energy harvested by sensor $i$ per unit of time. Note that different sensors have different rates $\lambda_k^i$ since various energy sources power the sensors. Further, the same energy harvester may also hold distinct energy supply capabilities, such as different sizes of panels in solar-powered equipment. In addition, the energy harvesting rate may change over time depending on location, weather conditions, and other factors [34]. Overall, $p_{k,q}^i$ implies the probability of $q$ units of energy harvested by sensor $i$ at time instant $k$.

**Remark 3:** According to the statement in [23], energy harvesting models are classified into two categories. One is deterministic models, and the other is stochastic models containing time-correlated and time-uncorrelated models. The Poisson process (7) is applied, considering that the majority of energy harvesting modules are composed of independent harvesting submodules. The energy arrival of each submodule can be modeled by a Bernoulli process when utilizing time-uncorrelated resources, such as wind and water energy. Thus, this enables the net energy harvesting to be demonstrated as a Binomial process. A Poisson process is approached at the limit when the number of submodules becomes large [35]. Moreover, the estimation method in this article can be extended to other types of energy harvesting models, such as the Markov model in [21] for time-correlated resources (e.g., solar energy), provided that the probability distribution of the amount of energy harvested by the sensor is known.

Denote the augmented state $\bar{x}_k \triangleq [\bar{x}_k^T \gamma_k^T]^T$. Combining (1), (2), and (4) leads to

$$\bar{x}_{k+1} = \bar{f}_k(\bar{x}_k) + B_k \bar{x}_k + \bar{w}_k$$

(8)

$$\bar{z}_k = C_k^i \bar{x}_k + v_k$$

(9)

with

$$\bar{f}_k(\bar{x}_k) = [f_k^T(x_k), 0]^T, \quad B_k = \begin{bmatrix} 0 & B_k \\ 0 & D_k \end{bmatrix}$$

$$\bar{w}_k = [w_k^T 0]^T, \quad C_k^i = [C_k 0].$$

Note that $\bar{w}_k$ is zero mean with covariance $\bar{Q}_k = \text{diag}\{Q_k, 0\}$.

Based on the above communication scenario limited by energy harvesting, the estimator for sensor $i$ is

$$\hat{x}_{k+1}^i = \hat{f}_k(\hat{x}_k) + \hat{B}_k \hat{x}_k$$

(10)

$$\hat{x}_{k+1}^i = \bar{x}_{k+1}^i + K_{k+1}^i (\bar{z}_{k+1}^i - \hat{x}_{k+1}^i)$$

$$+ G_{k+1}^i \sum_{j \in \mathcal{N}_k} (\beta_j^i \bar{z}_{k+1}^j + \mu_j^i \bar{x}_{k+1}^j)$$

(11)

where $\hat{x}_{k+1}^i \in \mathbb{R}^{n_{x,i} + n_{z,i}}$ is the one-step prediction, $\bar{z}_{k+1}^i \in \mathbb{R}^{n_{z,i} + n_{z,i}}$ is the estimate for the state $\bar{x}_k$ with the initial value $\bar{x}_0 = [\bar{z}_0^T 0]^T$, and $\hat{z}_{k+1}^i = C_k^i \hat{x}_{k+1}^i$. In addition, $K_{k+1}$ and $G_{k+1}$, respectively, represent the estimator gain and the consensus gain to be designed, and $\mu_j^i = \mathbb{E}\{\beta_j^i\}$ is used to reflect the information of whether the data $\bar{x}_{k+1}^j$ of sensor $j$ is successfully transmitted due to the unavailability of $\beta_{j+1}^i$, $j \in \mathcal{N}_k^i$ to sensor $i$.

**Remark 4:** The probability distribution of the indicator variable $\beta_k^i$ in (11) is based on the evolution of the energy level in (6). Hence, $\beta_k^i$ is related to $\beta_k^i$ for $k \neq i$. This feature is the most significant difference of (11) as compared with the existing missing measurement or intermittent observation models (e.g., [7], [8]), which rely on the Bernoulli random variable with a known probability distribution.

**Remark 5:** A novel energy-based transmission scenario is modeled to reflect the engineering reality. Specifically, the transmission decision in (5) depends on the number of neighbors of the sensor, that is, more neighbors require more energy consumption. This feature differs from [22], [23], and [33], in which communication can occur only when the energy level exceeds zero.

For simplicity, denote the prediction error $\eta_{k+1}^i = \bar{x}_{k+1} - \hat{x}_{k+1}^i$, and the estimation error $\hat{\eta}_k^i = \bar{x}_k - \hat{x}_k^i$. Correspondingly, the prediction and estimation error covariances are denoted by $P_k^i = \mathbb{E}\{\bar{x}_k^i \mid (\bar{x}_0^i, \hat{\eta}_0^i)^T\}$ and $P_{k+1}^i = \mathbb{E}\{(\hat{\eta}_{k+1}^i, \hat{\eta}_{k+1}^i)^T\}$, respectively. With (8) and (10) in hand, the prediction error is written as

$$\eta_{k+1}^i = \bar{f}_k(\hat{x}_k) - \bar{f}_k(\hat{x}_k^i) + B_k \hat{\eta}_k^i + \bar{w}_k.$$
\[ \Theta_{k,k+1} = \mathbb{E}\{G_{k+1}^i \sum_{j \in \mathcal{N}}: \beta_{k+1}^j \gamma_{k+1}^j \tilde{x}_k^T(\beta_{k+1}^j - \mu_{k+1}^j)(G_{k+1}^j)^T \} \]

Note that \( P_{k+1|k} \) in (14) includes the unknown schemes resulting from the system nonlinearity. Moreover, due to the intricate network communication and the uncertainty of energy harvesting, it is challenging to calculate the cross-covariance between different sensors involved in \( P_{k+1}^i \) in (15). Hence, the exact value for the error covariance \( P_{k+1}^i \) is unavailable. Alternatively, one can obtain an upper bound of \( P_{k+1}^i \) and then design the energy-based estimation parameters to minimize this bound. It is also desirable to reflect the effects of faults, system nonlinearity, energy harvesting constraints, and interconnections among sensors on this upper bound.

### III. MAIN RESULTS

#### A. Calculation of Sensor Transmission Probability

Owing to the intermittent data transmission caused by the randomness of energy harvesting, it is necessary to analyze the sensor transmission probability. First, the probability distribution of the energy level \( E_k^i \) of the system is discussed. Denote \( \phi_k^i = [Pr(E_k^i = 0) \quad Pr(E_k^i = 1) \quad \cdots \quad Pr(E_k^i = \tilde{E}_i)]^T \), and \( \bar{\phi}_k^i = [Pr(E_k^i = 0) \quad Pr(E_k^i = 1) \quad \cdots \quad Pr(E_k^i = \tilde{E}_i)]^T \).

**Lemma 1:** The probability distribution \( \phi_k^i \) of the energy level \( E_k^i \) of sensor \( i \) is established recursively as follows:

\[
\phi_{k+1}^i = \varepsilon^i + \Phi_k^i \bar{\phi}_k^i
\]

\[
\Phi_k^i = \begin{bmatrix}
0 & \cdots & 0 & E_0^i & \cdots & E_{\tilde{E}_i-1}^i & \tilde{E}_i
\end{bmatrix}
\]

and matrix \( \Phi_k^i \) is provided in (*), shown at the bottom of the page.

**Proof:** See Appendix A.

Notice that the initial condition of the energy level indicates the value of \( \phi_0^i \) in (16). Additionally, the structure of \( \varepsilon^i \) and \( \Phi_k^i \) guarantees that the sum of the probabilities of all possible energy levels at any time instant \( k \) is one. Moreover, when \( n^i = 1 \) and \( p_{k,0}^i = p_q \) for all \( i \in \mathcal{N}, k \geq 0 \), the probability matrix \( \Phi_k^i \) in (*) is reduced to the one used in [22]. Hence, the matrix \( \Phi_k^i \) in (*) is the generalized form of the probability matrix of harvesting.

From Lemma 1, the transmission probability of sensor \( i \) is

\[
\mu_k^i = \begin{bmatrix}
0 & \cdots & 0 & 1 & \cdots & E_{n^i-1}^i & \tilde{E}_i
\end{bmatrix}
\]

#### B. Estimator Design

In this section, an upper bound of \( P_{k+1}^i \) is obtained recursively and minimized.

**Theorem 1:** Given the positive scalars \( a_k^i, b_k^i, c_1^i, c_2^i, c_3^i \), and \( \xi_k^i \). Denote \( \xi_1^i = 1 + c_1^i k + c_2^i k^2, \xi_2^i = 1 + (c_2^i)^2 + c_3^i k, \) and \( \xi_3^i = 1 + (c_2^i)^2 + (c_3^i)^2 k \). The recursions are satisfied as follows:

\[
P_{k+1}^i = \xi_1^i k^i k_{k+1}^i P_{k+1|k}^i m_{k+1}^i (M_{k+1}^i)^T + K_{k+1}^i R_{k+1}^i (K_{k+1}^i)^T
\]

\[
+ \xi_2^i k_{k+1}^i k_{k+1}^i G_{k+1}^i \sum_{j \in \mathcal{N}} \mu_{j+1}^i P_{k+1|k}^j (G_{k+1}^j)^T
\]

\[
+ \xi_3^i k_{k+1}^i k_{k+1}^i G_{k+1}^i \sum_{j \in \mathcal{N}} [\mu_{j+1}^i - (\mu_{j+1}^i)^2] \bar{X}_{k+1}^i (G_{k+1}^i)^T
\]

with the initial condition \( P_0^i \geq P_0^i \), where \( \bar{X}_k^i = \text{diag} \{ A_k^i, 0 \} \), \( I = [I_{n^i}, 0_{n^i \times n^i}] \), and

\[
\bar{X}_k^i = (1 + b_k^i k_{k+1}^i) P_{k+1|k}^i + (1 + (b_k^i k_{k+1}^i)^2) k_{k+1}^i (k_{k+1}^i)^T
\]

Then, \( P_{k+1}^i \) shown in (19) is an upper bound of \( P_{k+1}^i \), that is, \( P_{k+1}^i \leq P_{k+1}^i \). Moreover, \( P_{k+1}^i \) is minimized with the gains

\[
K_{k+1}^i = \mathcal{O}_{k+1}^i (C_{k+1}^i)^T (C_{k+1}^i) + \mathcal{O}_{k+1}^i (C_{k+1}^i)^T + R_{k+1}^i)^{-1}
\]

\[
G_{k+1}^i = \xi_1^i k_{k+1}^i \sum_{j \in \mathcal{N}_k^i} \mu_{j+1}^i (I - K_{k+1}^i C_{k+1}^i) P_{k+1|k}^j (S_{k+1}^i)^{-1}
\]

\[
S_{k+1}^i = \xi_1^i k_{k+1}^i \left( \sum_{j \in \mathcal{N}_k^i} \mu_{j+1}^i \right)^2 P_{k+1|k}^i + \xi_2^i k_{k+1}^i \sum_{j \in \mathcal{N}_k^i} \mu_{j+1}^i
\]

\[
+ \bar{P}_{k+1|k}^i + \xi_3^i k_{k+1}^i \sum_{j \in \mathcal{N}_k^i} [\mu_{j+1}^i - (\mu_{j+1}^i)^2] \bar{X}_{k+1}^i
\]

\[
\mathcal{O}_{k+1}^i = \xi_1^i k_{k+1}^i \left( \sum_{j \in \mathcal{N}_k^i} \mu_{j+1}^i \right)^2 P_{k+1|k}^i (S_{k+1}^i)^{-1}
\]

**Proof:** See Appendix B.

**Remark 6:** The effects of the considered aspects (i.e., faults, system nonlinearity, and energy-constrained transmission protocol) on the upper bound are described here. Specifically, for \( P_{k+1|k}^i \) in (18) and \( P_{k+1}^i \) in (19), \( \mathcal{O}_k^i \) characterizes the dynamics of the fault signal, \( \alpha_k \) and \( \bar{A}_k \) refer to the system nonlinearity, and \( \mu_{k+1}^i \) corresponds to the energy-constrained
transmission protocol. Thus, this allows the effects from the considered aspects to be reflected in a unified distributed estimation framework.

Remark 7: The gains $K_{k+1}$ and $G_{i,k+1}$ in (20) are devised to minimize the upper bound $P_{k+1}^i$ in (19) such that the upper bound $P_{k+1}$ is as tight as possible to the original $P_{k+1}$. Moreover, it can be seen in (18) and (19) that the minimal upper bound $P_{k+1}^i$ is affected by the tuning parameters $a_i, b_i, c_{i,k+1}, d_{i,k+1}, e_{i,k+1}, f_{i,k+1}$ and $g_{i,k+1}$. Therefore, one can tune these parameters appropriately to cause the upper bound $P_{k+1}^i$ to be tighter to the original $P_{k+1}$. In addition, the matrix multiplication and inversion operations mainly lead to the computational complexity of the estimator design. Recalling $x_k \in \mathbb{R}^{n_r}$ and $\gamma_k \in \mathbb{R}^n$, one can conclude that the computational complexity on each sensor node is $O((n_x + n_r)^3)$.

C. Convergence Analysis

This section discusses the convergence performance of the proposed estimator.

**Corollary 1:** Assume that there exists real positive scalars $\xi_1, \bar{\alpha}, \bar{\beta}$ and $\bar{\eta}$ such that

$$\dim_{F_k} \leq \xi_1, \quad \alpha_k \leq \bar{\alpha}, \quad Q_k \leq \bar{\eta}I, \quad \|\bar{A}_k + \bar{B}_k\| \leq \bar{\beta}.$$ 

Then, if $\delta = \xi_1(1 + \bar{\alpha})^2 + \bar{\eta}
\leq 1$, one gets

$$\lim_{k \to \infty} P_k \leq \frac{p}{(1 - \delta)}I,$$

where $p = \max(\lambda_{\max}(P_k), \xi_1 \bar{\eta})$.

**Proof:** See Appendix C.

**Corollary 1** gives a sufficient condition to guarantee the convergence of the proposed estimator. However, it is observed that this sufficient condition imposes the requirements on the system parameters $\alpha_k$ and $\bar{A}_k + \bar{B}_k$. Roughly speaking, there must be $\alpha_k \leq \bar{\alpha} < 1$ and $\|\bar{A}_k + \bar{B}_k\| \leq \bar{\beta} < 1$ to ensure $\delta < 1$. Thus, the two aspects of convergence can be summarized as follows. 1) For some systems, $\alpha_k$ and $\bar{A}_k \in \mathbb{R}^{n_r}$ satisfying (3) cannot allow the convergence condition in Corollary 1 to hold and 2) $\bar{B}_k$ in (8) containing $\bar{B}_k$ and $\bar{D}_k$ determines the size of the fault signal. Hence, the proposed convergence condition is suitable for small fault signals satisfying $\|\bar{A}_k + \bar{B}_k\| \leq \bar{\beta} < 1$. In general, the convergence condition in Corollary 1 is somewhat impractical. Thus, this condition is relaxed. An alternative target performance of the estimator is to guarantee the exponential boundedness of the estimation error $\eta_k$ in the mean square. This performance can be easily achieved by applying the method in [36], and only the boundedness of the system parameters is required.

IV. SIMULATION

In this section, the proposed estimator is applied to an industrial isothermal continuous stirred tank reactor (CSTR) system [37], [38], where cyclopentadiene ($A$) reacts to form cyclopentenol ($B$); $A \rightarrow B$. The physical structure of the CSTR system is shown in Fig. 2, where only the educt $A$ in low concentration $C_{A0}$ flows into the reactor. The reactor outflow is a mixture of the desired product $B$ and educt $A$. Here, $C_A$ and $C_B$ denote the concentrations of $A$ and $B$ within the reactor, respectively.

![Fig. 2: Schematic diagram of the CSTR system.](image)

From [37], [38], the balance equations of the reactor are

$$\begin{align*}
\frac{dC_A}{dt} &= F \left( C_{A0} - C_A - k_1 C_A - k_2 C_B \right) \\
\frac{dC_B}{dt} &= -\frac{F}{V} (C_{A0} - C_A - k_3 C_A - k_2 C_B)
\end{align*}$$

(21)

where $F$ is the volume flow, $V$ is the reactor volume, and $k_1, k_2$, and $k_3$ are the rate coefficients. Let the subscript $s$ correspond to the steady-state value. Denote $x = [x_1 \ x_2]^T = [C_A - C_{As} \ C_B - C_{Bs}]^T$, $u = [F/V - F_s/V \ C_{A0} - C_{As} - C_{Bs}]^T$, and $z = x_2 = [C_B - C_{Bs}]$. The linearized state space model for (21) at the steady-state solution is $\dot{z} = \Xi x + Tu$ and $z = Hx + Lw$ with

$$\Xi = \begin{bmatrix} -\frac{F_s}{V} - k_1 & 0 \\ k_1 & -\frac{F_s}{V} - k_2 \end{bmatrix}, \quad H = \begin{bmatrix} 0 & 1 \end{bmatrix}, \quad L = \begin{bmatrix} 0 & 0 \end{bmatrix}.$$ 

Choose the parameters given in [37], [38] at the operating point 300.14 °C as follows, $k_1 = k_2 = 0.625$ min$^{-1}$, $k_3 = 0.0865$ l/(mol·min)$^{-1}$, $F_s/V = 0.3138$ min$^{-1}$, $C_{As} = 1.235$ mol/l, $C_{A0,s} = 5.1$ mol/l, and $C_{Bs} = 0.9$ mol/l.

In practice, nonlinearities are ubiquitous due to the networked-induced phenomena and are often described as additive nonlinear perturbations. Furthermore, faults are another common phenomenon that may degrade the performance of engineering systems. Hence, based on the CSTR model in [37], [38], system nonlinearity and faults are also considered. In this case, using a zero-order holder with a sampling period 0.4 min, a discretized model generated from MATLAB is derived in the form of (1) and (4) with the following parameters

$$f_k(x_k) = \begin{bmatrix} 0.6307 x_{1,k} + 0.05 \sin(x_{1,k}) \\ 0.1646 x_{1,k} + 0.6869 x_{2,k} + 0.05 \sin(x_{2,k}) \end{bmatrix}$$

$$B_k = \begin{bmatrix} 0.55 \\ 0.55 \end{bmatrix}^T, \quad C_k = \begin{bmatrix} 0 & 1 \end{bmatrix}.$$ 

(22)

It can be immediately seen from (22) that

$$A_k = \begin{bmatrix} 0.6307 & 0 \\ 0.1646 & 0.6869 \end{bmatrix}$$

and $\alpha_k = 0.05$ allow the nonlinear function $f_k(x_k)$ to satisfy condition (3). The dynamics of the fault signal are expressed in the form of (2) with $D_k = 1.95 \cos(0.5k)$. Then, the proposed distributed estimator is utilized to estimate the concentrations $C_A$ and $C_B$. A network with
three cooperative sensors monitors the reactor to improve the estimation performance. The network topology is represented in Fig. 2. Moreover, each sensor $i$ is equipped with an energy harvester, and the amount of energy harvested $H_{ik}^1$ follows the Poisson process in (7) with harvesting rate $\lambda_k^1$. Here, we choose $\lambda_k^1 = \lambda = 0.8$ for $i \in \mathcal{N}$. According to the transmission rule (5), a sensor transmits data to its neighbors if its current energy level is sufficient.

The involved scalars are chosen as $a_k^i = b_k^i = c_{1,k}^i = c_{2,k}^i = 2$. Set $Q_k = 0.001I$, $R_k^1 = 0.001$, $E_{i}^0 = 1$, and $E_{i}^3 = 4$. By using Theorem 1, the parameters of the estimator (10), (11) are obtained recursively.

The corresponding simulation results are presented in Figs. 3-6. Specifically, the actual state $x_{i,k}$ ($l = 1, 2$) and its estimates $\hat{x}_{i,k}$ for $i \in \mathcal{N}$ are shown in Fig. 3. The fault signal $\gamma_k$ and its estimates $\hat{\gamma}_k$ are described in Fig. 4. The above simulation results demonstrate that the proposed distributed estimator can reasonably estimate the concentrations and the faults of the CSTR system. Define the mean square error (MSE) of sensor $i$ at time instant $k$ as

$$\text{MSE}_k^i = \frac{1}{T} \sum_{m=1}^{T} \left[ \sum_{l=1}^{2} (x_{i,k}^m - \hat{x}_{i,k}^m)^2 + (\gamma_k^m - \hat{\gamma}_k^m)^2 \right]$$

where the superscript $m$ stands for the $m$th run. All MSE results are derived from $T = 300$ independent Monte Carlo runs. The trace of the upper bound $\bar{P}_k^1$ and the MSE in Eq. (23) are shown in Fig. 5. Note that the results of sensors 2 and 3 are similar to those of sensor 1 and therefore omitted. It can be seen from Fig. 5 that $\text{MSE}_k^1$ always lies below $\bar{P}_k^1$, which verifies that $\bar{P}_k^1$ obtained in (19) is indeed an upper bound of the error covariance matrix. Fig. 6 shows the energy level $E_k^1$ and the transmission decision $\beta_k^1$ for sensor $i = 1, 2$. Comparing (a) and (b) in Fig. 6, it is concluded that sensor $2$, which has a smaller number of neighbors, is more prone to transmit data. This demonstration is consistent with the designed transmission rule (5). In addition, it can be intuitively observed that sensors will not perform data transmission when energy is not available.

Subsequently, Fig. 7 reveals the effect of the energy harvesting rate $\lambda$ on the estimation performance. Notice from Fig. 7 that the estimation performance gradually improves (trace of $\bar{P}_k^1$ decreases) as the value of the energy harvesting rate $\lambda$ enlarges. This phenomenon is because an energy harvester with a higher harvesting rate can replenish more energy for the sensor. More power tends to induce frequent data transmission, which is conducive to estimation performance.

Finally, the fault is changed to be constant to illustrate the relationship between faults and energy. The selection of other parameters is the same as mentioned above. For comparison purposes, the fault is $\gamma_k = 0.1$ before $k = 50$ and is increased to $\gamma_k = 0.5$ after $k = 50$. The fault MSE of sensor 1 defined by $(1/T) \sum_{m=1}^{T} (\gamma_k^m - \hat{\gamma}_k^m)^2$ with $T = 300$ Monte Carlo runs is used as an evaluation indicator. The comparison of fault MSEs under different energy harvesting rates is depicted in Fig. 8. One can notice from Fig. 8 that for $\gamma_k = 0.1$, the fault estimation accuracy using the same harvesting rate $\lambda$ always remains similar. When fault $\gamma_k = 0.5$, the estimation accuracy of $\lambda = 0.2$ and $\lambda = 0.4$ gets worse over time. However, $\lambda = 0.6$ and $\lambda = 0.8$ still have stable performance. Thus, as the faults increase, more energy is needed to allow better fault estimation performance. In other words, sufficient energy can overcome the negative impact of the faults on the system to a certain extent. For instance, in real applications, it is necessary to consume more power to maintain the routine manipulation of legacy systems with many faults.
V. CONCLUSIONS

A distributed estimator for simultaneously estimating the state and the fault has been proposed for a category of nonlinear time-varying systems over sensor networks under energy harvesting constraints. A recursive calculation of the sensor transmission probability has been provided by employing the probability distribution of the energy amount. An upper bound of the estimation error covariance has first been established considering faults, nonlinearity, and energy-related transmission protocol. This bound has been minimized by selecting suitable filter parameters. The convergence analysis of the proposed estimator has been discussed. The article has ended with a numerical example to exhibit the validity of the main results. Future directions worthy of research include applying the proposed distributed energy-based estimator to 1) sensor networks that allow energy sharing among sensors and 2) time-varying sensor networks.

APPENDIX A

PROOF OF LEMMA 1

The calculation of the probability of the energy level \( E_{k+1}^i = m, 0 \leq m \leq E_i^i - 1 \) is divided into the following cases where \( q = \min\{m, E_i^i - n^i\} \).

**Case 1:** For \( 0 \leq m < n^i \), there are two situations:

**Case 1.1:** Recalling (7), if \( n^i \leq E_i^i \), one can obtain

\[
\Pr(E_{k+1}^i = m) = \sum_{s=0}^{m} p_{k,m-s}^i \Pr(E_k^i = s) + \sum_{s=0}^{q} p_{k,m-s}^i \Pr(E_k^i = n^i + s).
\]

**Case 1.2:** Otherwise, that is, for \( n^i > E_i^i \),

\[
\Pr(E_{k+1}^i = m) = \sum_{s=0}^{m} p_{k,m-s}^i \Pr(E_k^i = s).
\]

**Case 2:** For \( m \geq n^i \), one has

\[
\Pr(E_{k+1}^i = m) = \sum_{s=0}^{n^i-1} p_{k,m-s}^i \Pr(E_k^i = s) + \sum_{s=0}^{q} p_{k,m-s}^i \Pr(E_k^i = n^i + s).
\]

On the other hand, the probability of energy level \( E_{k+1}^i = E_i^i \) is calculated by \( \Pr(E_{k+1}^i = E_i^i) = 1 - \sum_{s=0}^{n^i-1} \Pr(E_k^i = s) \).

In summary, the recursive process of the probability distribution of \( E_{k+1}^i \) is shown in (16).

APPENDIX B

PROOF OF THEOREM 1

**Lemma 2 ([32]):** For \( x, y \in \mathbb{R}^a \) and a positive scalar \( \kappa \), one can get \( xy^T + yx^T \leq \kappa x^T x + \kappa^{-1} y^T y \).

Then, the proof of Theorem 1 is provided. In the light of (3) and (14), one has

\[
P_{k+1}^i \leq (1 + a_k^i) \alpha_k^i \left[ (I \alpha_k^i)^T I + (1 + (a_k^i)^{-1}) \right] \times (\hat{A}_k + \hat{B}_k) P_{k+1}^i (\hat{A}_k + \hat{B}_k)^T + \hat{Q}_k.
\]

By virtue of Lemma 2, it can be derived that in (15),

\[
\Omega_{k+1}^i \leq \frac{1}{2} G_{k+1}^i \left\{ \sum_{j \in \mathcal{N}_i} \sum_{\eta_j} \left[ (\eta_j^i + 1) \bar{\eta}_j^i (\eta_j^i + 1)^T \right] \right\} (G_{k+1}^i)^T
\]

Hence, similar to (25), we obtain

\[
\Gamma_{k+1}^i \leq n^i G_{k+1}^i \sum_{j \in \mathcal{N}_i} \left[ (\mu_j^i + 2) \bar{X}_{k+1}^i (G_{k+1}^i)^T \right] \triangleq \tilde{\Gamma}_{k+1}^i.
\]

Applying Lemma 2 to the terms \( \Theta_{k+1}^i (\Theta_{k+1}^i)^T \), \( -\Theta_{k+1}^i (\Theta_{k+1}^i)^T \), and \( -\Theta_{k+1}^i (\Theta_{k+1}^i)^T \) in (15), and then combining (25) and (26), we have

\[
P_{k+1}^i \leq \xi_{k+1}^i M_{k+1}^i + P_{k+1}^i d^i \left( M_{k+1}^i + K_{k+1}^i R_{k+1}^i (K_{k+1}^i)^T \right)
\]

It follows from (14), (15), (18), (19), (24), and (27) that \( P_{k+1}^i \leq \tilde{P}_{k+1}^i \). Next, by solving \( \partial \tilde{P}(P_{k+1}^i) / \partial K_{k+1}^i = 0 \) and \( \partial \tilde{P}(P_{k+1}^i) / \partial G_{k+1}^i = 0 \), one gets the optimal gains \( K_{k+1}^i \) and \( G_{k+1}^i \) in (20) such that \( \tilde{P}_{k+1}^i \) is minimized.

APPENDIX C

PROOF OF COROLLARY 1

Substituting \( K_{k+1}^i \) and \( G_{k+1}^i \) in (20) into (19), the equivalent expression of (19) is obtained as follows:

\[
P_{k+1}^i = \mathcal{O}_{k+1}^i - \mathcal{O}_{k+1}^i (C_{k+1}^i)^T
\]

\[
\times \left[ C_{k+1}^i \mathcal{O}_{k+1}^i (C_{k+1}^i)^T \right]^T + R_{k+1}^i - \bar{C}_{k+1}^i \mathcal{O}_{k+1}^i
\]

\[
= \left[ (\mathcal{O}_{k+1}^i)^{-1} + (C_{k+1}^i)^T \right] (R_{k+1}^i)^{-1} C_{k+1}^i \leq \mathcal{O}_{k+1}^i
\]

where the last equation is from the matrix inversion lemma.

Denote \( \Lambda_{k+1}^i = \xi_{k+1}^i (\sum_{j \in \mathcal{N}_i} \mu_j^i \bar{X}_{k+1}^i)^2 \). Recalling the definition of \( S_{k+1}^i \), we have \( (S_{k+1}^i)^{-1} < (\Lambda_{k+1}^i)^{-1} \), and hence \( 0 < \Lambda_{k+1}^i (S_{k+1}^i)^{-1} < I \). There must exist a \( 0 < \tau < 1 \) such that \( \Lambda_{k+1}^i (S_{k+1}^i)^{-1} > \tau I \). Then, \( \mathcal{O}_{k+1}^i \) can be rearranged as \( \mathcal{O}_{k+1}^i = \xi_{k+1}^i \left[ (I - \Lambda_{k+1}^i (S_{k+1}^i)^{-1}) \right] \mathcal{O}_{k+1}^i \) becomes

\[
\tilde{P}_{k+1}^i < \xi_{k+1}^i \mathcal{O}_{k+1}^i.
\]
Inserting (18) into (29) yields
\[
P_{k+1}^i < \xi_i^k (1 + \alpha_i^k \lambda_0 \text{tr} \{ \hat{P}_{k+1}^i \hat{H}_k \hat{E}_k \}) (1 + (\alpha_i^k)^{-1}) \times (\bar{A}_i + \bar{B}_i \hat{P}_{k}^i (\bar{A}_i + \bar{B}_i)^T + \xi_i^k + i_{k+1} Q_i).
\]  
(30)

Based on (30), we use mathematical induction to complete the following derivations. Denote \( p = \max \{ \lambda_{\max} (P_0^i), \xi_1^i, I \} \). Due to \( P_0^i > 0 \), we have \( P_k^i \leq \lambda_{\max} (P_0^i) I \leq p I \). Combining with (30), one gets \( P_{k+1}^i < \delta p I + p I \). Then, suppose that \( P_{k-1}^i < p^{\sum_{m=0}^{k-1} s_m n} I \), it is derived that \( P_k^i < \delta p \sum_{m=0}^{k-1} s_m n + p I = p^{k+1} \). Hence, one has \( \lim_{k \to \infty} P_k^i < (p/1-\delta) I \) if \( \delta < 1 \).

REFERENCES


