A novel iterative identification based on the optimized topology for common state monitoring in Wireless Sensor Networks

Zhenyu Lu [a], Ning Wang [a] and Chenguang Yang [a]*

[a] *Bristol Robotics Laboratory, University of the west of England, Bristol, UK;*

*Email: cyang@ieee.org

A novel iterative identification based on the optimized topology for common state monitoring in Wireless Sensor Networks

Power consumption and data redundancy of wireless sensor networks (WSN) are widely considered for a distributed state monitoring network. For reducing the energy consumption and data amount, we propose a topology optimization and an iterative parameter identification method for estimating the common model factors in WSN. The former method optimizes the decentralized topology such that all the leaf nodes in a community connect to the head node directly. A circle topology is built to enable the remote leaf nodes to link to the head node through two adjoining relay nodes to reduce the whole communication distance and power consumption. Based on the optimized topology, an iterative identification method is proposed to minimize the information capacity by transmitting the processed results instead of raw data to reduce the data amount for calculation and storage. Then, we prove the consensus and convergence of the proposed identification method. Finally, two simulations verify the effectiveness of the proposed method and the comparative results present the data reduction for the on-board calculation, communication, and storage in the practical use of WSN.

Keywords: Wireless sensor networks (WSN); Power consumption; Iterative parameter identification; Convergence;

1. Introduction

WSN is consisted of large numbers of low-cost and low-energy sensors, with ability of collecting observations, on-board processing and wireless communication (Akyildiz and Su et al., 2002; Ruiz-Garcia, Lunadei, Barreiro & Robla 2009). In recent decades, the technique of WSN gains much attention in a wide range of areas including robot locating (Chen, Lu, & Peng, et al., 2019; Li, Wang, & Wang, et al.,2020; Zong, Ji, & Yu, et al., 2020), and environment, construction and agriculture monitoring (Yang, Huang, Zhang, & Hua, 2016; Pakzad, Rocha, & Yu, 2011; Zhang, Yu, Song & Wang, 2013). High communication power consumption, data redundancy and no global identification (ID) for sensor nodes are main challenges for WSN (Dutta, Gupta, & Das 2012, Moschitta &

Neri 2014, Rawat, Singh, Chaouchi, & Bonnin 2014). The recent work of Rawat, Singh, Chaouchi, & Bonnin (2014) presented that the power consumption for receiving circuitry is greater than the consumptions for transmitting circuitry and the baseband digital signal processing also consumes a lot of energy. Therefore, reducing data amount for delivery and calculation is helpful to reduce energy consumption.

For reducing the data amount for communication, some scholars studied the data compression technology (Alsheikh, Lin, Niyato, & Tan, 2016), data fusion (Collotta, Pau & Bobovich, 2017; Lin, Chen, & Varshney 2005; Liu, Zhu & Anjum et al., 2020) and data filtering (Bashir, Lim, Hussain, & Park, 2011) to balance the demands of calculation and energy consumption in WSN. Parameter identification is widely used in WSN-based structure healthy monitoring (SHM) systems (Sim, & Spencer, 2009; Nagayama, & Spencer 2007) and helps to reduce data amount by delivering the processed results instead of raw data. A typical application is healthy monitoring of large bridges (like Golden Gate Bridge) using a WSN (Dorvash, & Pakzad, 2013). The vibration properties of the structure i.e. natural frequencies, damping ratios and mode shapes are identified using the measured response of the system by WSN to the environmental or forced excitations.

WSN-based parameters identification is topology-related. In Kim et al., 2007; Chintalapudi et al., 2006, the topologies of WSN are divided into two groups: centralized group and independent group. The centralized group has poor scalability and high energy consumptions (Liu, Cao & Lai, et al., 2011). The independent group has a powerful data processing ability for each sensor node and the data processing outputs are sent back to the base station. The following researches improved the independent group and clustered sensor nodes into several communities (in Figure 1) and named it decentralized network topology (Cho, Park & Sim, 2015). In each community, there is only one head node and

several leaf nodes linked to the head. The head node receives and sends processed outputs to the base station.

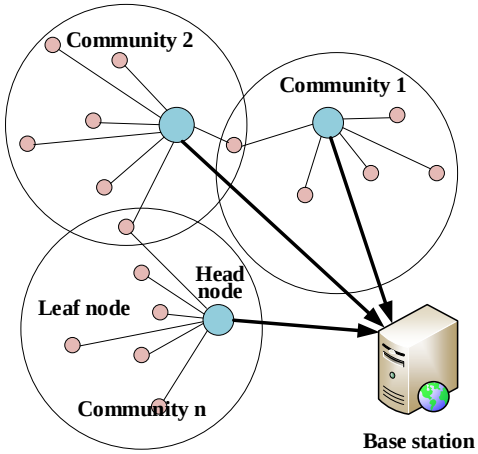


Figure 1. Decentralized network topology (the WSN is divided into several local sensor communities (circle), each of which contains one head node (blue node) and multiple leaf nodes (red node), and all the head nodes connect the central base station (computer))

For the identification, the data are also processed on the sensor board then sent to the head node or a high-level base station. Some researchers have proposed methods for the decentralized structure. Pakzad, Rocha, & Yu (2011) utilized the distributed modal identification approaches to estimate the parameters in regularized autoregressive models that eliminate the computation of correlation between signals measured at nodes far apart. Nagayama, & Spencer (2007) and Sim, & Spencer (2009) divided the WSN into several sub-networks with cluster heads with a hierarchical topology and proposed an interactive modal identification (IMID) method for structural health monitoring, which can reduce the burden in time and energy and optimize the allocation. Liu & Cao et al. (2011) pointed out that each sensor node in most methods performs modal analysis only based on its own measuring data, and input change or measurement noise can easily degrade the identified local modal parameters and the errors cannot be reduced in the assembling process at the central unit. Thus they divided the whole network into a number of single-hop clusters (a hop occurs when a packet is passed from one network node to the centre node), and a cluster head (CH) is designated in each cluster to perform intra-cluster modal analysis.

Cho, Park & Sim (2015) also utilized a stochastic subspace identification method to cluster leaf nodes for the decentralized identification.

Compared with centralized mode, the decentralized identification reduces the data amount for communication with the head node, but it still can't be used for the multi-hop WSN, which is consisted of several nodes, and some nodes are communicated through multiple hops, though its coverage area is larger than radio range of single nodes (Pešović, Mohorko, Benkič, & Čučej (2010)). Therefore, the nodes far away from the head will use other nodes as relays. Even if the relay nodes can help the remote nodes to deliver data to the head, it will cost a lot of energy and time. For the problem, we proposed a topology optimization method to transform the topology that all leaf nodes directly connecting to the head node into a multi-hop network. Then each remote leaf node just communicates with the two nearest neighbouring nodes. A novel iterative identification method is proposed based on the new topology that each leaf node can calculate, based on own measurements and delivered information from its neighbours. The delivered data is decreased to the minimum amount of model factors for identification. Additionally, the computational complexity and energy consumption also decrease by avoiding inverse calculation of large matrices, which is shown in Table 2. Considering the head node cannot get the first-hand information from every leaf node directly and the processed results may cause estimation errors and bring uncertainties to the final conclusions of the head node, we prove that all the leaf nodes will acquire the global unbiased conclusions.

The rest of this paper is organized as follows: Section 2 presents a topology optimization and an iterative parameters identification method. We prove the consensus and convergence of identified results for every leaf node and relay node based on several assumptions. Section 3 verifies the effectiveness of the proposed methods by simulations,

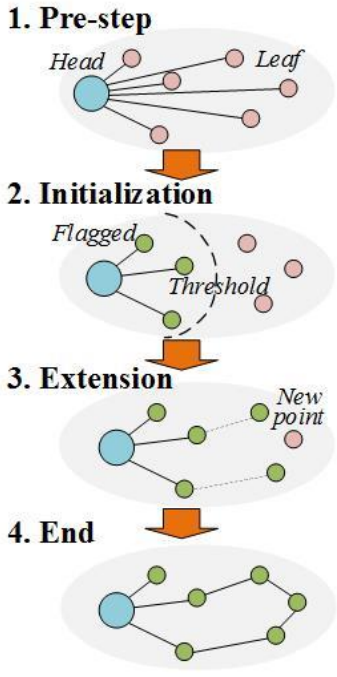
and compares the data amount for communication, calculation and storage with those in other methods. Section 4 concludes this paper in the end.

2. Topology optimization and iterative identification method

2.1 Topology optimization algorithm and discussion

Nagayama and Spencer (2007) proposed a NExT-based data processing approach, which divides the sensor network into several local sensor communities, and every community consists of a cluster-head node and several leaf nodes. The head node sends the processed data to its connected leaf nodes as references for correction, and every leaf node calculates correlation function based on the reference and own measurements. But this method only suits the single-hop WSN given that all the leaf nodes connect to the head node, even for remote ones. In this section, we will further optimize the connection topology based on the results of NExT-based method to reduce the distance of remote nodes linked to the head one.

```
1  # define Flag          1;
2  # define NoFlag        0;
3  # define Threshold     H;  // H is a number
4  Function Connection topology
5  Start Node 1, Node m Flag, Node 2,3,...m-1 NoFlag;
6        Calculate the distance of Node 2,3,...m-1
7        If the distance of k and the head<Threshold then
8              Node k (k∈{2,3,...m-1}) Flag
9        End if
10 Node 1, Node m is the current point;
11 If existing node with Flag
12       For (current point: i=1; i<2;i++)
13             Calculate the distance of Node 2,3,...m-1with NoFlag to
14             current point
15             If (the distance of Node n is the minimum)&Node n NoFlag
16             then
17                   Node n Flag;
18                   Node n is current point;
19             End if
20       End for
21 End if
```



Algorithm 1. Topology optimization algorithm and illustrations

Algorithm 1 presents the algorithm as well as the topology optimization process. First, in the pre-work step, all the leaf nodes connected to the head node by the clustering algorithm. After setting a threshold (line 3, initialization step), we initialize the flagged nodes whose distance are closer than the threshold and select two of them as the 1^{st} and m^{th} nodes for the new topology. The two nodes are marked with flags and the current nodes are used for searching relay nodes to extend the communication channels (line 10). The extended nodes are selected as new searching nodes till all the remote leaf nodes are flagged (Extension and End steps on the right of Algorithm 1, and lines 11 to 21 on the left of Algorithm 1) to realize the connectivity of each node is 2.

Seen from the last step named 'End' in Algorithm 1, the optimized topology is a combination of several single-hop connections and a multi-hop connection. The remote leaf node does not need to make a long-distance communication with the head, hence reducing energy consumption. Compared with the decentralized topology in Cho, Park & Sim (2015), the new topology enables the head node communicates with its neighbouring nodes to reduce power consumption for covering the whole community. But, the relay nodes should process the received information and send the results to neighbours such that their calculation burdens are increased. In Bashir, & Lim et al. 2011, the information is delivered with a feedback message mechanism with confirmation in a circle across the head node and all the leaf and delay nodes. Time delay is always a core issue for network system control (Peng, Qiao, & Xu, 2002; Wang & Qiao 2002). For the problems of long time delays and the increasing calculation burdens of relay nodes caused by the optimized topology, we propose a new iterative identification method.

2.2 Iterative parameter identification method.

In WSN, the data is collected with low installation and maintenance cost at high spatial

and temporal resolution (Pakzada, Rochab & Yu, 2011), but with measuring biases and noise caused by calibration and installation errors etc. Researchers used Autoregressive with Exogenous model (ARX) (Dorvash, Pakzad, & Cheng, 2013), Autoregressive Moving Average with Exogenous (ARMAX) models, output-error (OE), Box–Jenkins (BJ) (Green, Nadimi, & Blanes-Vidal, 2009) and multivariate autoregressive (AR) model (Pakzada, Rochab & Yu, 2011) etc. to fit the measurements of the environment structure's response. As each node processes the data for monitoring the same object, the model of identification is the same for every node. In this paper, we chose the ARMAX model to express the dynamics of the *ith* node as

$$\begin{aligned} y_i(k) &= -\sum_{m=1}^{p} a_m y_i(k-m) + \sum_{n=0}^{q} b_n x_i(k-j) + \sigma_i(k) \\ &= \boldsymbol{\varphi}_i(k)^T \boldsymbol{\theta} + \sigma_i(k) \end{aligned}, \qquad (1)$$

where $y_i(k)$ and $x_i(k)$ are output and input vectors at the kth sampling time, a_m and b_n are coefficients, and $\sigma_i(k)$ is measuring noise, and p and q are system orders for the autoregressive and exogenous parts. $\boldsymbol{\varphi}_i(k) = [y_i(k-1), \ldots, y_i(k-p), x_i(k), \ldots, x_i(k-q)]^T$ is an information vector ,and $\boldsymbol{\theta} = [a_1, a_2, \ldots, a_p, b_0, b_1, \ldots, b_q]^T$ is a parameter vector for identification.

Set $\hat{\boldsymbol{\theta}}(k)$ as the estimation of $\boldsymbol{\theta}$ in time interval $[kh, kh+h]$, and h is the signal sampling interval for every sensor node. Several assumptions are proposed as follows.

- **Assumption1**. All the sensors are isomorphic and their models share the common parameters $\boldsymbol{\theta}$, and the interval for signal sampling, delivering and processing is h.
- **Assumption2**. Signals of all the nodes will be processed asynchronously, but the procedures will be completed within the interval h till the beginning of the next

circulation. Set $\tau_{i,i+1}$ as the communication time delay of node i and j and τ_i^c as calculation time of node i, then in a community with m leaf nodes, we have

$$\sum_{i=1}^{m}\tau_i^c+\sum_{i=1}^{m-1}\tau_{i,i+1}+\tau_{m,1}<h.$$

- **Assumption3**. Within time interval $[kh,kh+h]$, with the increase of inputting data amount, the value of $\hat{\boldsymbol{\theta}}_i(k)$ at time k is closer to the real value $\boldsymbol{\theta}$ than the previous nodes $\hat{\boldsymbol{\theta}}_j(l),\ j=1,2,...,m,l=1,2,...k-1$ in the former sampling time interval.
- **Assumption4**. Exist constants $0<\alpha\le\beta<\infty$ and $N\ge n$ satisfying the condition

$$\alpha\boldsymbol{I}\le\frac{1}{N}\sum_{j=0}^{N-1}\boldsymbol{\varphi}_i(k+j)\boldsymbol{\varphi}_i(k+j)^T\le\beta\boldsymbol{I}\ .$$

- **Assumption5**. The noise signal $\sigma_i(t)$ is an orthogonal factor, and the expectation of $\sigma_i(k)\sigma_j(k)$ is bounded with Δ^2

$$\left|\sigma_i(k)\sigma_j(k)\right|=\begin{cases}0 & i\ne j\\ \sigma^2\le\Delta^2 & i=j\end{cases},i=1,2,...,m.$$

Remark 1: Assumptions 1 and 2 are about time delays and proposed as the basic demands for the network. The two assumptions ensure the information for all the sensor nodes in a community are measured, processed and delivered within a sampling interval. Assumption 3 to 5 are basic theoretical bases for the parameter identification, which is introduced in (Ding, Shi, & Chen, 2006) and Ding, F. (2012) etc. Assumption 3 shows the calculations based on the previous results are closer to the real values in statistic with the increase of data amount. Assumptions 4 and 5 present the input signals and noises are bounded and energy-limited physically.

The traditional parameter identification methods like in Ding, Shi, & Chen, 2006 are calculated based on own measurements as

$$\begin{cases} \hat{\boldsymbol{\theta}}_i(k) = \hat{\boldsymbol{\theta}}_i(k-1) + \mathbf{L}_i(k)[y_i(k) - \boldsymbol{\varphi}_i^T(k)\hat{\boldsymbol{\theta}}_i(k-1)] \\ \mathbf{L}_i(k) = \mathbf{P}_i(k-1)\boldsymbol{\varphi}_1^T(k)[\mathbf{I} + \boldsymbol{\varphi}_i^T(k)\mathbf{P}_i(k-1)\boldsymbol{\varphi}_i(k)]^{-1} \\ \mathbf{P}_i(k) = [\mathbf{I} - \mathbf{L}_i(k)\boldsymbol{\varphi}_i^T(k)]\mathbf{P}_i(k-1) \end{cases} \quad (2)$$

The information cannot be shared and delivered between different sensor nodes. Based on the optimized connecting topology, we propose a new iterative identification in Figure 2.

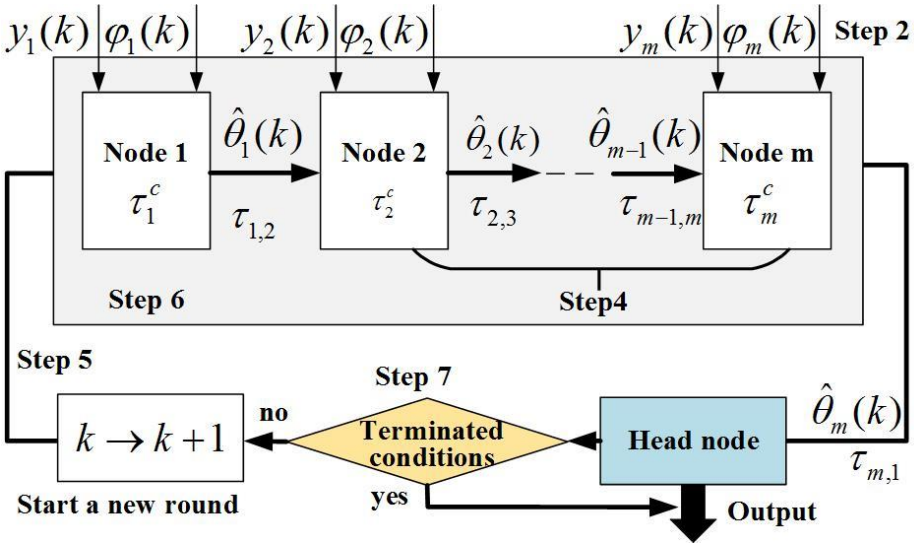


Figure 2. Flow chart of the iterative parameter identification method.

The sensor nodes are located in the circle-topology route and transmit information in sequence from node 1 to node m and the head node. Every leaf node will collect $\boldsymbol{\varphi}_i(k)$ and $y_i(k)$ from the environment and receive the processed information $\hat{\boldsymbol{\theta}}_{i-1}(k)$ received (or $\hat{\boldsymbol{\theta}}_m(k-1)$ for node 1) from the former node $i-1$. The information is processed using the following eqs. (3) to (5) to get $\hat{\boldsymbol{\theta}}_i(k)$. After waiting $\tau_{i,i+1}$, the ith node will receive $\hat{\boldsymbol{\theta}}_i(k)$ and continue the calculation and data delivery till node m. The last node m sends $\hat{\boldsymbol{\theta}}_m(k)$ as the end of the kth cycle to the head node, and the head node will make an

average calculation using $\hat{\boldsymbol{\theta}}_m(k)$ and the identifications of other independent leaf nodes and then send the final results to the base station and node 1 to start the next cycle. The procedures and mathematical expressions are presented as follows:

(1) Initialize all the parameters and take $\hat{\boldsymbol{\theta}}_i(0)=\mathbf{1}_{1\times n}/p_0$, $\mathbf{P}_i(k)$ is a covariance matrix, $\mathbf{P}_i(0)=p_0\mathbf{I}_n$, $i=1,2,...,m$. p_0 is a large enough number, and $\mathbf{L}(k)$ is a gain matrix and $\mathbf{I}_n$ is a unit matrix.

(2) Collect data $y_i(k)$ and $\boldsymbol{\varphi}_i(k)$ in every leaf node respectively.

(3) Initialize the parameter $\hat{\boldsymbol{\theta}}_1(1)$ of node 1 by

$$\begin{cases}\hat{\boldsymbol{\theta}}_1(1)=\hat{\boldsymbol{\theta}}_1(0)+\mathbf{L}_1(1)[y_1(1)-\boldsymbol{\varphi}_1^{\,T}(1)\hat{\boldsymbol{\theta}}_1(0)]\\ \mathbf{L}_1(1)=\mathbf{P}_1(0)\boldsymbol{\varphi}_1^{\,T}(1)[\mathbf{I}+\boldsymbol{\varphi}_1^{\,T}(1)\mathbf{P}_1(0)\boldsymbol{\varphi}_1(1)]^{-1}\\ \mathbf{P}_1(1)=[\mathbf{I}-\mathbf{L}_1(1)\boldsymbol{\varphi}_1^{\,T}(1)]\mathbf{P}_1(0)\end{cases}. \tag{3}$$

Then $\hat{\boldsymbol{\theta}}_1(1)$ is delivered to node 2.

(4) According to Assumption 3, within sampling interval $[kh,kh+h]$, $\hat{\boldsymbol{\theta}}_{i-1}(k)$ is closer to the real value $\boldsymbol{\theta}$ than $\hat{\boldsymbol{\theta}}_i(k-1)$ acquired in $[kh-h,kh]$. Compared with (2), we prefer to use $\hat{\boldsymbol{\theta}}_{i-1}(k)$ to replace $\hat{\boldsymbol{\theta}}_i(k-1)$ in (2) for the calculation of $\hat{\boldsymbol{\theta}}_i(k)$, and $\hat{\boldsymbol{\theta}}_i(k)$ from node 2 to node m in Figure 2 are calculated by (4) as

$$\begin{cases}\hat{\boldsymbol{\theta}}_i(k)=\hat{\boldsymbol{\theta}}_{i-1}(k)+\mathbf{L}_i(k)[y_i(k)-\boldsymbol{\varphi}_i^{T}(k)\hat{\boldsymbol{\theta}}_{i-1}(k)]\\ \mathbf{L}_i(k)=\mathbf{P}_i(k-1)\boldsymbol{\varphi}_i(k)/[1+\boldsymbol{\varphi}_i^{T}(k)\mathbf{P}_i(k-1)\boldsymbol{\varphi}_i(k)]\\ \mathbf{P}_i(k)=[\mathbf{I}-\mathbf{L}_i(k)\boldsymbol{\varphi}_i^{T}(k)]\mathbf{P}_i(k-1),\quad i=2,3,...,m\end{cases}. \tag{4}$$

(5) Node m sends $\hat{\boldsymbol{\theta}}_m(k)$ to the head node, and the head node transmits $\hat{\boldsymbol{\theta}}_m(k)$ to node 1 at the same time.

(6) Node 1 estimates $\hat{\boldsymbol{\theta}}_1(k+1)$ based on $\hat{\boldsymbol{\theta}}_m(k)$ by (5) to start the next circulation.

$$\begin{cases} \hat{\boldsymbol{\theta}}_1(k+1) = \hat{\boldsymbol{\theta}}_m(k) + \mathbf{L}_1(k+1)[y_1(k+1) - \boldsymbol{\varphi}_1^T(k+1)\hat{\boldsymbol{\theta}}_m(k)] \\ \mathbf{L}_1(k+1) = \mathbf{P}_1(k)\boldsymbol{\varphi}_1(k+1)/[1+\boldsymbol{\varphi}_1^T(k+1)\mathbf{P}_1(k)\boldsymbol{\varphi}_1(k+1)] \\ \mathbf{P}_1(k+1) = [\mathbf{I} - \mathbf{L}_1(k+1)\boldsymbol{\varphi}_1^T(k+1)]\mathbf{P}_1(k) \end{cases} . \quad (5)$$

(7) Continue the calculating process till the terminated conditions.

Remark 2: Seen from (4) and (5), the estimated result $\hat{\boldsymbol{\theta}}_i(k)$ is calculated based on $\hat{\boldsymbol{\theta}}_{i-1}(k)$ of the previous node, and $\hat{\boldsymbol{\theta}}_{i-1}(k)$ is based on $\hat{\boldsymbol{\theta}}_{i-2}(k)$, which means any identified results are the common effect of all the sensor nodes linked by the circle topology.

2.3 Consensus and convergence proofs

In this section, we will prove the proposed identification method achieves the same consistent and converging effect as the centralized and decentralized modes. Consensus means the estimations for all the sensor nodes keep the same, and convergence means the estimated parameters converge to the real values, eliminating measuring noise and biases. The proof contains two steps. The first is to prove consensus for all the parameters. If we can prove any leaf node can converge to the real value, all the estimations will converge.

Submitting (5) into (4), we can get

$$\begin{aligned}\hat{\boldsymbol{\theta}}_2(k) &= \hat{\boldsymbol{\theta}}_1(k) + \mathbf{L}_2(k)\left[y_2(k) - \boldsymbol{\varphi}_2^T(k)\hat{\boldsymbol{\theta}}_1(k)\right] \\ &= \hat{\boldsymbol{\theta}}_m(k-1) + \mathbf{L}_1(k)\left[y_1(k) - \boldsymbol{\varphi}_1^T(k)\hat{\boldsymbol{\theta}}_m(k-1)\right] \\ &\quad + \mathbf{L}_2(k)\left[y_2(k) - \boldsymbol{\varphi}_2^T(k)\hat{\boldsymbol{\theta}}_1(k)\right] \\ &= \hat{\boldsymbol{\theta}}_m(k-1) + \mathbf{L}_1(k)\left[y_1(k) - \boldsymbol{\varphi}_1^T(k)\hat{\boldsymbol{\theta}}_m(k-1)\right] \\ &\quad + \mathbf{L}_2(k)\left[y_2(k) - \boldsymbol{\varphi}_2^T(k)\hat{\boldsymbol{\theta}}_m(k-1)\right] \\ &\quad - \mathbf{L}_2(k)\boldsymbol{\varphi}_2^T(k)\left[\hat{\boldsymbol{\theta}}_1(k) - \hat{\boldsymbol{\theta}}_m(k-1)\right]\end{aligned}. \tag{6}$$

Then, a recursive conclusion can get

$$\hat{\boldsymbol{\theta}}_m(k) = \hat{\boldsymbol{\theta}}_m(k-1) + \mathbf{L}(k)^T \Delta\mathbf{Y}(k) - \sum_{i=2}^{m} \mathbf{L}_i(k)\boldsymbol{\varphi}_i^T(k)\left[\hat{\boldsymbol{\theta}}_{i-1}(k) - \hat{\boldsymbol{\theta}}_m(k-1)\right], \tag{7}$$

where $\mathbf{L}(t) = [\mathbf{L}_1(k), ..., \mathbf{L}_m(k)]^T$, $\mathbf{Y}(t) = [y_1(k), ..., y_m(k)]^T$, $\Delta\mathbf{Y}(k) = \mathbf{Y}(k) - \boldsymbol{\Psi}^T(k)\hat{\boldsymbol{\theta}}_m(k-1)$, and $\boldsymbol{\Psi}(t) = [\boldsymbol{\varphi}_1(k), ..., \boldsymbol{\varphi}_m(k)]^T$.

Set $\tilde{\boldsymbol{\theta}}_m(k-1) = \boldsymbol{\theta} - \hat{\boldsymbol{\theta}}_m(k-1)$, then the *ith* item of $\Delta\mathbf{Y}(t)$ is

$$\begin{aligned}\Delta y_i(k) &= y_i(k) - \boldsymbol{\varphi}_i^T(k)\hat{\boldsymbol{\theta}}_m(k-1) \\ &= \boldsymbol{\varphi}_i^T(k)\tilde{\boldsymbol{\theta}}_m(k-1) + \sigma_i(k)\end{aligned}. \tag{8}$$

Eq. (7) can be expressed as

$$\begin{aligned}&\hat{\boldsymbol{\theta}}_m(k) - \hat{\boldsymbol{\theta}}_m(k-1) + \sum_{i=2}^{m} \mathbf{L}_i(k)\boldsymbol{\varphi}_i^T(k)[\hat{\boldsymbol{\theta}}_{i-1}(k) - \hat{\boldsymbol{\theta}}_m(k-1)] \\ &= \mathbf{L}(k)^T[\boldsymbol{\Psi}(k)\tilde{\boldsymbol{\theta}}_m(k-1) + \boldsymbol{\Delta}(k)]\end{aligned}, \tag{9}$$

where $\boldsymbol{\Delta}(k) = [\sigma_1(k), \sigma_2(k), \cdots, \sigma_m(k)]^T$. If the right equation of (9) converges to 0, then the left equation equals to 0 too. We will introduce the following two lemmas.

Lemma 1. For any item $\boldsymbol{\varphi}_i^T(k)\boldsymbol{\varphi}_j(k)$, it satisfies

$$\boldsymbol{\varphi}_i^T(k)\boldsymbol{\varphi}_j(k) \le ((\beta - \alpha)k + \alpha)\mathbf{I}, \quad i, j = 1, 2, 3, ..., m.$$

The proof of lemma is presented in appendix.

Define two variables $e_i(k) = \boldsymbol{\varphi}_i^T(k)\tilde{\boldsymbol{\theta}}_m(k-1) + \sigma_i(k)$, $\eta_i(k) = \boldsymbol{\varphi}_i^T(k)\tilde{\boldsymbol{\theta}}_m(k) + \sigma_i(k)$, then $\tilde{\boldsymbol{\theta}}_m(k)$ can be expressed by $\tilde{\boldsymbol{\theta}}_m(k-1)$ and the error items $e_i(k)$ and $\eta_i(k)$ as

$$\begin{aligned}\tilde{\boldsymbol{\theta}}_m(k) &= \tilde{\boldsymbol{\theta}}_m(k-1) - \sum_{i=1}^{m} k_{mi}(k)\mathbf{P}_i(k)\boldsymbol{\varphi}_i(k)e_i(k) \\ &= \tilde{\boldsymbol{\theta}}_m(k-1) - \sum_{i=1}^{m} k_{mi}(k)\mathbf{P}_i(k-1)\boldsymbol{\varphi}_i(k)\eta_i(k)\end{aligned}, \tag{10}$$

where $k_{mi}(k) = \begin{cases} \prod_{j=i+1}^{m}(\mathbf{I} - \mathbf{L}_j(k)\boldsymbol{\varphi}_j^T(k)) = \prod_{j=i+1}^{m}\mathbf{P}_j(k)\mathbf{P}_j(k-1)^{-1}, i \neq m \\ 1 \qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\quad , i = m \end{cases}$.

Define $\mathbf{T}(t)$ as

$$\begin{aligned}\mathbf{T}(k) &= \tilde{\boldsymbol{\theta}}_m(k-1)^T \boldsymbol{\Psi}(k)\boldsymbol{\Psi}(k)^T \tilde{\boldsymbol{\theta}}_m(k-1) \\ &= \sum_{i=1}^{m} \tilde{\boldsymbol{\theta}}_m(k-1)^T \boldsymbol{\varphi}_i(k)\boldsymbol{\varphi}_i^T(k)\tilde{\boldsymbol{\theta}}_m(k-1)\end{aligned}. \tag{11}$$

Lemma 2. Based on the previous assumptions and Lemma 1, we can get when $k \to \infty$, the expectation of $\mathbf{T}(k)$ satisfies

$$\left|\mathbf{T}(k)\right| < 2\left(\frac{\beta}{\alpha} + 1\right)m^2\Delta^2. \tag{12}$$

The proof of lemma 2 is presented in appendix.

Set $\Gamma(k) = \boldsymbol{\Psi}^T(k)\tilde{\boldsymbol{\theta}}_m(k-1)$, then $\mathbf{T}(k)$ is expressed as $\mathbf{T}(k) = \Gamma^T(k)\Gamma(k)$. Using (12), we knows that the expectation of $\Gamma(k)$ is bounded. So expectation of the right item $\boldsymbol{\Psi}^T(k)\tilde{\boldsymbol{\theta}}_m(k-1) + \boldsymbol{\Delta}(k)$ in (9) is also bounded. When $k \to \infty$, $\mathbf{P}_i(k)$ will decrease to zero, then any vector $\mathbf{L}_i(k) = \boldsymbol{\varphi}_i^T(k)\mathbf{P}_i(k)$ in $\mathbf{L}(k)$ will decrease to zero gradually. Namely, the

right formula in (9) converges to zero finally, so does the left. The left equation of (9) can be written as

$$\lim_{k\to\infty}\left[\hat{\boldsymbol{\theta}}_m(k)-\hat{\boldsymbol{\theta}}_m(k-1)+\sum_{i=2}^{m}\mathbf{L}_i(k)\boldsymbol{\varphi}_i^T(k)[\hat{\boldsymbol{\theta}}_{i-1}(k)-\hat{\boldsymbol{\theta}}_m(k-1)]\right]=0. \qquad (13)$$

Similar to the simplifying process of $\hat{\boldsymbol{\theta}}_m(t)$, we can get

$$\begin{aligned}&\hat{\boldsymbol{\theta}}_1(k)-\hat{\boldsymbol{\theta}}_1(k-1)+\mathbf{L}_1(k)\boldsymbol{\varphi}_1^T(k)[\hat{\boldsymbol{\theta}}_m(k-1)-\hat{\boldsymbol{\theta}}_1(k-1)]+\\&\sum_{i=2}^{m-1}\mathbf{L}_{i+1}(k)\boldsymbol{\varphi}_{i+1}^T(k)[\hat{\boldsymbol{\theta}}_i(k-1)-\hat{\boldsymbol{\theta}}_1(k-1)]=0\end{aligned}. \qquad (14)$$

Successive substitution proves that for every node, the relationship of $\lim_{k\to\infty}\hat{\boldsymbol{\theta}}_i(k)$ and $\lim_{k\to\infty}\hat{\boldsymbol{\theta}}_i(k-1)$ can be expressed in a unique form as

$$\hat{\boldsymbol{\theta}}_i(k)=\hat{\boldsymbol{\theta}}_i(k-1)-\sum_{j\in N_t,\,j\neq i}\mathbf{L}_s(k)\boldsymbol{\varphi}_s^T(k)[\hat{\boldsymbol{\theta}}_j(k)-\hat{\boldsymbol{\theta}}_i(k-1)],\ i=1,2,...,m, \qquad (15)$$

where $N_t=1,...,m$ and $s=\begin{cases}j+1, & j=1,2,...,m-1\\1\ \ , & j=m\end{cases}$.

The consensus of $\lim_{k\to\infty}\hat{\boldsymbol{\theta}}_i(k), i=1,2,...,m$ can be proved by theorem 1.

Theorem 1. For a system

$$\begin{cases}\Phi_i=\phi_i(k+1)=\phi_i(k)+\mu_i(k) & i\in N_t\\ \mu_i(k)=-\sum_{j\neq i,\,j\in N_t}a_{ij}[\Phi_j-\phi_i(k)]\end{cases}, \qquad (16)$$

where $N_t=1,...,m$ and $0<a_{ij}<1/(m-1)$ is a random number . If $\forall K$, $\phi_i(k)$ converge to the stable state at time $k>K$, then all the state values of $\phi_i(k)$ will approach to a

constant simultaneously.

Proof of Theorem 1.

Following (16), we have

$$\phi_1(k+1)-\phi_1(k)=-\sum_{j=2}^{m}a_{1j}\left[\phi_j(k)-\phi_1(k)\right], \tag{17}$$

$$\phi_2(k+1)-\phi_2(k)=-\sum_{j\neq 2\in N_t}a_{2j}\left[\phi_j(k)-\phi_2(k)\right]-\sum_{j=2}^{m}a_{21}a_{1j}\left[\phi_j(k)-\phi_1(k)\right]. \tag{18}$$

It is not hard to get $\phi_K(k+1)-\phi_K(k), K\in N_t$ can be expressed by $\phi_i(k)-\phi_j(k)$ as

$$\phi_K(k+1)-\phi_K(k)=F(\phi_i(k)-\phi_j(k)), i, j\in N_t. \tag{19}$$

Setting $\begin{bmatrix} I & \cdots & 0 & \cdots & 0 \\ -a_{j1} & \cdots & \sum_{j\neq i\in N_i}a_{ji} & \cdots & -a_{jK} \\ 0 & \cdots & 0 & \cdots & I \end{bmatrix}=\Psi_j$, $\begin{bmatrix}\phi_1(k)\\ \phi_2(k)\\ \vdots \\ \phi_K(k)\end{bmatrix}=\mathrm{X}(k)$, we can express the relation of $\mathrm{X}(k)$ and $\mathrm{X}(k+1)$ as

$$\begin{aligned}&\mathrm{X}(k+1)=\\ &\left\{\begin{bmatrix} I & 0 & \cdots & 0 \\ -a_{21} & I & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ -a_{m1} & -a_{m2}(1+a_{32}a_{21}-a_{32}-a_{31}) & \cdots & I \end{bmatrix}+\prod_{j=1}^{m}\Psi_j\right\}\mathrm{X}(k).\\ &=\mathbf{\Pi}(k)\mathrm{X}(k)\end{aligned} \tag{20}$$

Obviously, matrix $\mathbf{\Pi}(k)$ is a full-rank time-varying matrix. If we set $\phi_i(k)=c$, $i\in N_t$ and take it to (19), then $\phi_i(k+1)=\phi_i(k)=constant$. So $\mathbf{\Pi}(k)$ always has the eigenvector $[1,1,...,1]^T\in R^{m\times 1}$. As the identification processes are converged, according to

Lemma 3 in Olfati-Saber, Fax, & Murray (2007), the parameters $\phi_i(k)$ will converge to a constant simultaneously. □

After proving the consensus for all the sensor nodes, we use the following Lemma 3 and Lemma 4 to prove the parameters $\hat{\boldsymbol{\theta}}_i(t), i=1,2,...,m$ to converge to $\boldsymbol{\theta}$.

Lemma 3. For a system $y_m(k)=\boldsymbol{\varphi}_m^T(k)\boldsymbol{\theta}+\sigma_m(k)$ in (1), satisfied all above assumptions , then the following inequality holds

$$\left|V(k)+S(k)\right| \le V(k-1)+S(k-1)+2m^2\Delta^2,$$

where $\left|V(k)+S(k)\right|$ represents the expectation of $V(k)+S(k)$.

$$V(k)=\sum_{i=1}^{m}\tilde{\boldsymbol{\theta}}_m^T(k)\mathbf{P}_i(k)^{-1}\tilde{\boldsymbol{\theta}}_m(k),$$

$$S(k)=2\sum_{j=1}^{k}\sum_{i=1}^{m}k_{m1}(j)u_i(j)v_i(j),\ u_i(k)=\tilde{\boldsymbol{\theta}}_m^T(k)\boldsymbol{\varphi}_i(k),$$

$$v_i(k)=\tilde{\boldsymbol{\theta}}_m^T(k)\mathbf{P}_i(k-1)\sum_{j=1}^{m}\mathbf{P}_j(k-1)^{-1}\boldsymbol{\varphi}_i(k)+\frac{1}{2}\tilde{\boldsymbol{\theta}}_m^T(k)\boldsymbol{\varphi}_i(k).$$

The proof of Lemma 3 is presented in Appendix.

Lemma 4. (Theorem 1 & Corollary 1, Ding & Chen, 2004) Suppose that all assumptions and Lemma 3 hold, and α, β and N mentioned in assumption 4 satisfy $\beta=\alpha N^{c_0}$, where c_0 is a constant. Then for $c>1$, the parameters estimation errors associated with proposed algorithm hold

$$\lim_{k\to\infty}\left\|\tilde{\boldsymbol{\theta}}_m(k)-\boldsymbol{\theta}\right\|^2=O\left(\frac{(\ln N)^c}{N}\right)\to 0$$

Though the measuring noises for each node are not zero-mean, the boundless and orthogonal property of the noise signals guarantee the convergence of calculations.

Remark 3: Lemmas 1 and 2 are proposed for proving Theorem 1 of consensus of the all the state values $\phi_i(k)$ ($\hat{\boldsymbol{\theta}}_i(k)$ in (15)). After confirming all the identified parameters will

converge to the same value, we will use Lemmas 3 and 4 to prove the convergence of any identified parameter $\hat{\boldsymbol{\theta}}_i(k)$ converging to $\boldsymbol{\theta}$. The final converging errors depend on the noise signal $\sigma_i(t)$ and its boundary Δ, which have similar performances to the traditional identification algorithms like (2). The following simulations present the convergence of the proposed method in comparison with ordinary identification methods.

3. Simulation and analysis

The topology optimization algorithm and the iterative identification method are proposed separately but affect each other during the simulation. Optimization topology is the basis for parameter identification and determines the signal transmitting sequence for the leaf nodes. So the first simulation is about topology optimization and selects the head node located at $[0.2, 0.2]$ and 11 leaf points randomly distributed within a unit square. The threshold value is set as 0.4 (as the black dash line shown), and the simulation results are presented in Figure 3.

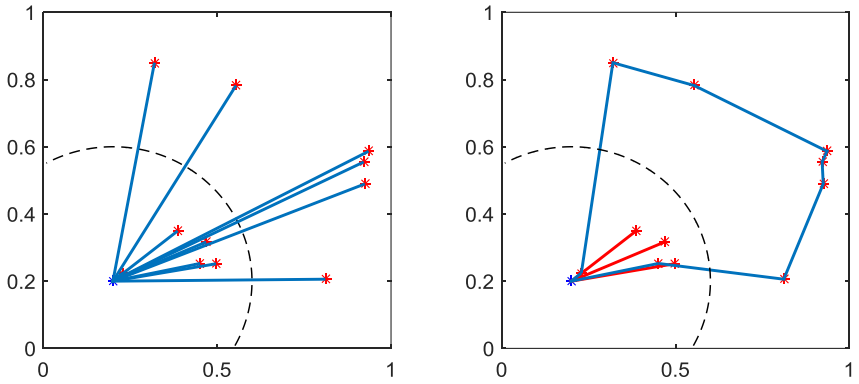


Figure 3. Original connecting topology and the optimized topology (a) Original connecting topology with multiple single-hops (b) Optimized connecting topology with some short-distance single-hops and a multi-hop link

There are two subfigures in Figure 3. Subfigure (a) shows the original connecting topology after point clustering such that all the nodes connect to the head node directly.

The whole length for signal communication of all the leaf nodes is 4.88. Subfigure (b) shows the optimized topology. The red lines represent the direct links between the leaf nodes and the head one and the blue lines represent the optimized circle topology. The whole length decreases to 3.86, which is about 80% of the original one. Compared with subfigure (a), a shorter communication length and a smaller covering range of the head node will save more energy for data transmission.

As the optimization process does not influence the direct connections of the leaf nodes nearby. Here, we consider the case of 8 leaf nodes located at the circle topology. The system function is

$$\begin{aligned} y_i(k) &= \boldsymbol{\varphi}_i^T(k)\boldsymbol{\theta} + \sigma_i(k) \\ &= x_i(k) - 1.15x_i(k-1) + 0.425x_i(k-2) + 0.55y_i(k-1) - 0.32y_i(k-2) + \sigma_i(k) \end{aligned},$$

where $\sigma_i(k)$ is the orthotropic noise with bounded variance $\Delta^2 = 1.0$, $\boldsymbol{\varphi}_i(k) = [x_i(k),$ $x_i(k-1), x_i(k-2), y_i(k-1), y_i(k-2)]^T$ is the information vector, and $\boldsymbol{\theta} = [a_0, a_1, a_2, b_1, b_2]^T$ $= [1, -1.15, 0.425, 0.55, 0.32]^T$ is the parameter vector.

We use the ordinary identification ((2) and the leaf nodes are independent) and iterative parameter identification algorithms ((3) to (5)) to estimate $\boldsymbol{\theta}$ based on the same conditions: initial values are $\hat{\boldsymbol{\theta}}_i(0) = 10^{-6} I^{1\times 5}$, $\mathbf{P}_i(0) = 10^6 I^{5\times 5}$ and sampling time interval is $h = 4.$ Time delays are set as random values within $[0.2, 0.5]$. The results are shown in Figure 4 (ordinary method) and Table 1 and Figure 5 (the proposed method).

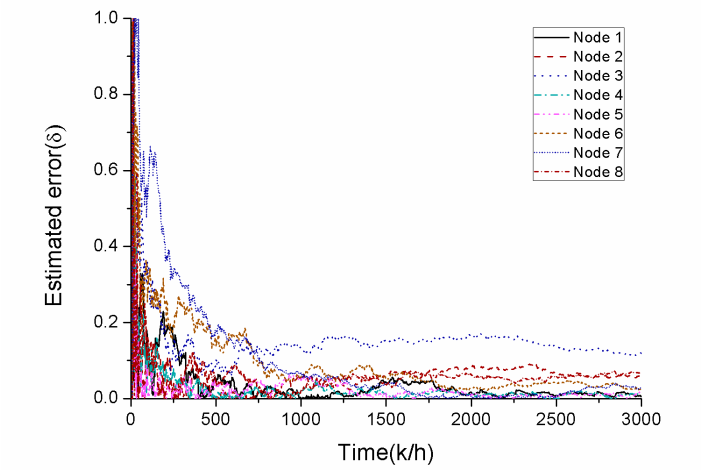


Figure 4. Estimation error rates δ vary with t in the normal identification method

Table 1. The parameter estimates ($\Delta^2 = 1.0$)

k	a_0	a_1	a_2	b_1	b_2	δ
100	0.912	-1.335	0.614	0.579	-0.412	8.52%
200	0.834	-1.299	0.477	0.592	-0.387	2.02%
400	0.881	-1.281	0.441	0.571	-0.324	1.70%
800	0.975	-1.198	0.470	0.564	-0.332	2.10%
1000	0.972	-1.183	0.473	0.569	-0.331	1.51%
2000	0.998	-1.143	0.467	0.558	-0.319	0.46%
3000	1.002	-1.150	0.448	0.551	-0.320	0.34%
True values	1.000	-1.150	0.425	0.550	-0.320	

* $\delta = \left\|\hat{\boldsymbol{\theta}}(k) - \boldsymbol{\theta}\right\| / \boldsymbol{\theta}$ **represents the relative parameter estimation error rate**

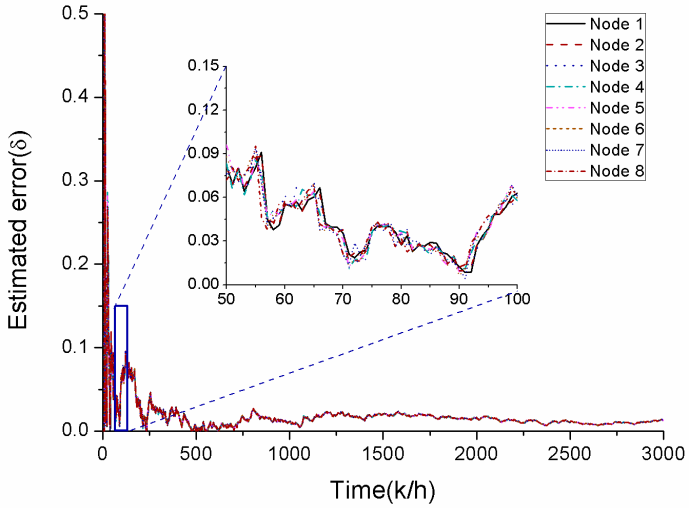


Figure 5. Estimation error rates δ vary with t in the iterative identification algorithm

Owing to the measuring noise, the estimation errors of different nodes converge to different values after about 1000 time iterations in Figure 4. Compared with Figure 4, the results in Figure 5 converge to the same states faster (less than 200 iterations) under the iterative calculation of all the nodes. The results verify the consensus and convergence of identification methods.

As power consumption is directly determined by data amount for communication, calculation and storage, we set the dimension of input data for a leaf node as N_i, the dimension of parameters as N_p ($N_p < N_i$), and the number of the leaf nodes as $K > 2$, then the data amount for communication, calculation and storage in different topologies are compared in Table 2 and further explained in Appendix D:

Table 2. Compare of data amount in different algorithms

		Centralized mode in Sim et al. 2009	Decentralized mode in Nagayama, et al. 2007	Proposed method in this paper
Communication	Leaf	KN_i	KN_p	KN_p
Calculation	Leaf	0	$K\left(o(N_i^3)\right)$	$K\left(o(N_i^3)\right)$
	Head	$o\left((KN_i)^3\right)$	KN_p	$2N_p$
Storage	Leaf	0	$K\left(N_i + N_p\right)$	$K\left(N_i + N_p\right)$
	Head	KN_i	KN_p	$2N_p$

Compared with centralized mode, the decentralized mode reduces communication data amount from KN_i to KN_p, and distributes the calculation burden and data storage of the head node to the leaf nodes. Because, for the leaf node, the calculation is based on own state vector that is different from the head node combines the state vectors into a large matrix for matrix inversion and multiplication, the calculation complexities for all the leaf and head nodes are reduced. The main improvement of the proposed method to the decentralized mode focuses on the topology. For the head node, the proposed method has fewer connections, thus that the data amounts for calculation and storage are reduced.

The proposed method also has advantages of shorter communication distance shown in simulation 1 and consistent processed results to eliminate the cognition difference of leaf sensor nodes. However, there are some limitations e. g. the optimized cycle topology and communication efficiency. Assumption 2 is also limited for a network with a larger cycle of leaf nodes or with small sampling time intervals. The proposed method suits the normal multi-hop topology as well, if the multi-poly lines can be seen as squashed circles that all the relay nodes and leaf nodes have two-way communications through the same channel.

Communication efficiency is hoped to improve by increasing the connectivity of sensor nodes. Referring Assumption 3 and the consensus theory, adding the topology connectivity will accelerate the speed of synchronization. Then the identified results can update the consistent state faster. We make a simulation for the multi-link case, whose topology is shown in figure 6:

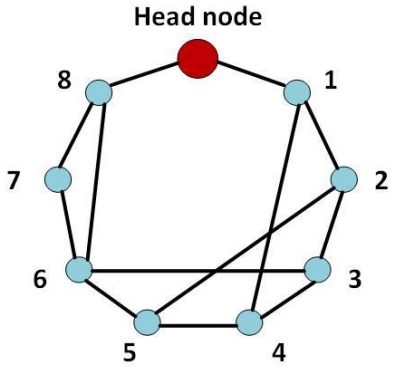


Figure 6. Multi-link topology

For every node, once the results are identified, the updated parameters will be transmitted to the neighbouring points. Similar to the assumptions proposed above, the calculation process is updated through the head node in a sampling period. The simulation results are presented as follows:

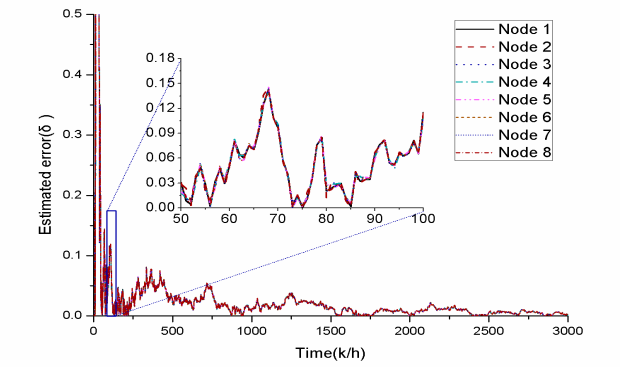


Figure 7. Parameter estimation error rates δ vary with t of multi-link topology

Compared with Figure 5, the identified results in Figure 7 approach to the consistent state faster with smaller differences. Time delay and packet dropout are also important for the WSN. In this paper, every node will update estimations based on the references transmitted from other nodes to reach to the asynchronous state. According to the consensus theory, the results can converge to the same state even if some packets are dropped. Whether the iterative identification method with multi-link topology can reduce the influence of packet dropout is worth further discussing.

4. Conclusion

For reducing power consumption and data redundancy in WSN, a topology optimization scheme and an iterative system identification method are proposed in this paper. Based on several assumptions, the consensus and convergence of the identified parameters are proved. Compared with the centralized and decentralized modes, the proposed method can reduce the data amount for communication, calculation and storage to some extent. Simulation results verify the effectiveness of the proposed method. The proposed iterative identification is also helpful for other distributed systems. As our current work is about human-inspired motion control (Yang, Chen, & Wang et al., 2018; Yang, Luo, & Liu et al.,2018; Yang, Wu, & Li, et al., 2017), the sensor nodes for measuring motions and EMG signals are expected to attached on human bodies for the large-group human behaviour monitoring and skill modelling.

References

Akyildiz, I. F., Su, W., Sankarasubramaniam, Y., & Cayirci, E. (2002). Wireless sensor networks: a survey. Computer networks, 38(4), 393-422.

Alsheikh, M. A., Lin, S., Niyato, D., & Tan, H.-P. (2016). Rate-distortion balanced data compression for wireless sensor networks. IEEE Sensors Journal, 16(12), 5072-5083.

Bashir, A. K., Lim, S.-J., Hussain, C. S., & Park, M.-S. (2011). Energy efficient in-network RFID data filtering scheme in wireless sensor networks. Sensors, 11(7), 7004-7021.

Chen, D., Lu, Q., Peng, D., Yin, K., Zhong, C., & Shi, T. (2019). Receding horizon control of mobile robots for locating unknown wireless sensor networks. Assembly Automation. 39 (3), 445-459

Chintalapudi, K., Paek, J., Gnawali, O., Fu, T. S., Dantu, K., Caffrey, J., Govindan, R., Johnson, E., & Masri, S. (2006). Structural damage detection and localization using NETSHM. 2006 5th International Conference on Information Processing in Sensor Networks, 475-482.

Cho, S., Park, J.-W., & Sim, S.-H. (2015). Decentralized system identification using stochastic subspace identification for wireless sensor networks. Sensors, 15(4), 8131-8145.

Collotta, M., Pau, G., & Bobovich, A. V. (2017). A fuzzy data fusion solution to enhance the QoS and the energy consumption in wireless sensor networks. Wireless Communications and Mobile Computing, 2017.

Cormen, T. H., Leiserson, C. E., Rivest, R. L., & Stein, C. (2009). Introduction to algorithms. MIT press.

Ding, F., & Chen, T. (2004). Combined parameter and output estimation of dual-rate systems using an auxiliary model. Automatica, 40(10), 1739-1748.

Ding, F., Shi, Y., & Chen, T. (2006). Performance analysis of estimation algorithms of nonstationary ARMA processes. IEEE Transactions on Signal Processing, 54(3), 1041-1053.

Ding, F. (2012). System identification. Part H, Coupled identification concept and methods. Journal of Nanjing University of Information Science and Technology, Natural Science Edition, 4(193), 212.

Dorvash, S., & Pakzad, S. (2013). Stochastic iterative modal identification algorithm and application in wireless sensor networks. Structural Control and Health Monitoring, 20(8), 1121-1137.

Dorvash, S., Pakzad, S. N., & Cheng, L. (2013). An iterative modal identification algorithm for structural health monitoring using wireless sensor networks. Earthquake Spectra, 29(2), 339-365.

Dutta, R., Gupta, S., & Das, M. K. (2012). Power consumption and maximizing network lifetime during communication of sensor node in WSN. Procedia Technology, 4, 158-162.

Fraleigh J. B. and Beauregard R. A., Linear Algebra. Addison-Wesley Publishing Company, 1995.

Green, O., Nadimi, E. S., Blanes-Vidal, V., Jørgensen, R. N., Storm, I. M. D., & Sørensen, C. G. (2009). Monitoring and modeling temperature variations inside silage stacks using novel wireless sensor networks. Computers and Electronics in Agriculture, 69(2), 149-157.

Kim, D.-S., Kashif, A., Ming, X., Kim, J.-H., & Park, M.-S. (2008). Energy efficient in-network phase RFID data filtering scheme. International Conference on Ubiquitous Intelligence and Computing, 311-322.

Kim, S., Pakzad, S., Culler, D., Demmel, J., Fenves, G., Glaser, S., & Turon, M. (2007). Health monitoring of civil infrastructures using wireless sensor networks. Proceedings of the 6th international conference on Information processing in sensor networks, 254-263.

Li, J., Wang, J., Wang, S., Peng, H., Wang, B., Qi, W., Zhang, L., & Su, H. (2020). Parallel structure of six wheel-legged robot trajectory tracking control with heavy payload under uncertain physical interaction. Assembly Automation. 40(5), 675-687.

Lin, Y., Chen, B., & Varshney, P. K. (2005). Decision fusion rules in multi-hop wireless sensor networks. IEEE Transactions on Aerospace and Electronic Systems, 41(2), 475-488.

Liu, B. (2016). Optimization of hierarchical data fusion in Wireless Sensor Networks. 2016 6th International Conference on Electronics Information and Emergency Communication.

Liu, X., Cao, J., Lai, S., Yang, C., Wu, H., & Xu, Y. L. (2011). Energy efficient clustering for WSN-based structural health monitoring. 2011 Proceedings IEEE INFOCOM.

Liu, X., Zhu, R., Anjum, A., Wang, J., Zhang, H., & Ma, M. (2020). Intelligent data fusion algorithm based on hybrid delay-aware adaptive clustering in wireless sensor networks. Future Generation Computer Systems, 104, 1-14.

Moschitta, A., & Neri, I. (2014). Power consumption assessment in wireless sensor networks (ICT-energy-concepts towards zero-power information and communication technology. IntechOpen.

Nagayama, T., & Spencer Jr, B. F. (2007). Structural health monitoring using smart sensors (No. 1940-9826

Olfati-Saber, R., Fax, J. A., & Murray, R. M. (2007). Consensus and cooperation in networked multi-agent systems. Proceedings of the IEEE, 95(1), 215-233.

Pakzad, S. N., Fenves, G. L., Kim, S., & Culler, D. E. (2008). Design and implementation of scalable wireless sensor network for structural monitoring. Journal of infrastructure systems, 14(1), 89-101.

Pakzad, S. N., Rocha, G. V., & Yu, B. (2011). Distributed modal identification using restricted auto regressive models. International Journal of Systems Science, 42(9), 1473-1489.

Pešović, U. M., Mohorko, J. J., Benkič, K., & Čučej, Ž. F. (2010). Single-hop vs. Multi-hop–Energy efficiency analysis in wireless sensor networks. In 18th Telecommunications Forum, TELFOR.

Peng, J., Qiao, H., & Xu, Z.-b. (2002). A new approach to stability of neural networks with time-varying delays. Neural networks, 15(1), 95-103.

Rawat, P., Singh, K. D., Chaouchi, H., & Bonnin, J. M. (2014). Wireless sensor networks: a survey on recent developments and potential synergies. The Journal of supercomputing, 68(1), 1-48.

Ruiz-Garcia, L., Lunadei, L., Barreiro, P., & Robla, I. (2009). A review of wireless sensor technologies and applications in agriculture and food industry: state of the art and current trends. Sensors, 9(6), 4728-4750.

Sim, S.-H., & Spencer Jr, B. F. (2009). Decentralized strategies for monitoring structures using wireless smart sensor networks No. 1940-9826

Wang, Z., & Qiao, H. (2002). Robust filtering for bilinear uncertain stochastic discrete-time systems. IEEE Transactions on Signal Processing, 50(3), 560-567.

Wang, Z., Qiao, H., & Burnham, K. J. (2002). On stabilization of bilinear uncertain time-delay stochastic systems with Markovian jumping parameters. IEEE Transactions on Automatic control, 47(4), 640-646.

Yang, C.-x., Huang, L.-y., Zhang, H., & Hua, W. (2016). Multirate parallel distributed compensation of a cluster in wireless sensor and actor networks. International Journal of Systems Science, 47(1), 1-13.

Yang, C., Chen, C., Wang, N., Ju, Z., Fu, J., & Wang, M. (2018). Biologically inspired motion modeling and neural control for robot learning from demonstrations. IEEE Transactions on Cognitive and Developmental Systems, 11(2), 281-291.

Yang, C., Luo, J., Liu, C., Li, M., & Dai, S.-L. (2018). Haptics electromyography perception and learning enhanced intelligence for teleoperated robot. IEEE Transactions on Automation Science and Engineering, 16(4), 1512-1521.

Yang, C., Wu, H., Li, Z., He, W., Wang, N., & Su, C.-Y. (2017). Mind control of a robotic arm with visual fusion technology. IEEE Transactions on Industrial Informatics, 14(9), 3822-3830.

Zhang, D., Yu, L., Song, H., & Wang, Q.-G. (2013). Distributed H∞ filtering for sensor networks with switching topology. International Journal of Systems Science, 44(11), 2104-2118.

Zong, C., Ji, Z., Yu, J., & Yu, H. (2020). An angle-changeable tracked robot with human-robot interaction in unstructured environments. Assembly Automation. 40 (4), 565-575.

Appendix

Appendix A. The proof of Lemma 1

Proof. Using Assumption 4, we have

$$\begin{cases} 2\boldsymbol{\varphi}_i^T(k)\boldsymbol{\varphi}_j(k) \le \boldsymbol{\varphi}_i^T(k)\boldsymbol{\varphi}_i(k) + \boldsymbol{\varphi}_j^T(k)\boldsymbol{\varphi}_j(k) \\ \boldsymbol{\varphi}_i^T(k)\boldsymbol{\varphi}_i(k) \le (\beta k - \alpha(k-1))\mathbf{I} \\ \boldsymbol{\varphi}_j^T(k)\boldsymbol{\varphi}_j(k) \le (\beta k - \alpha(k-1))\mathbf{I} \end{cases}. \tag{21}$$

Repeated substitution shows that

$$\boldsymbol{\varphi}_i^T(k)\boldsymbol{\varphi}_j(k) \le (\beta k - \alpha(k-1))\mathbf{I} = ((\beta-\alpha)k+\alpha)\mathbf{I}.$$

Appendix B The proof of Lemma 2

Proof. To simplify the proving process, let $\tilde{\boldsymbol{\theta}}(k) = \tilde{\boldsymbol{\theta}}_m(k)$ and substitute (10) into $\mathbf{T}(t)$:

$$\begin{aligned}\mathbf{T}(k) &= \sum_{i=1}^{m}\left\{\left[\tilde{\boldsymbol{\theta}}(k) + \sum_{j=1}^{m} k_{mj}(k)\mathbf{P}_j(k)\boldsymbol{\varphi}_j(k)e_j(k)\right]^T \mathbf{P}_i(k)^{-1}\left[\tilde{\boldsymbol{\theta}}(k) + \sum_{j=1}^{m} k_{mj}(k)\mathbf{P}_j(k)\boldsymbol{\varphi}_j(k)e_j(k)\right]\right. \\ &\qquad \left. - \tilde{\boldsymbol{\theta}}(k-1)^T\mathbf{P}_i(k-1)^{-1}\tilde{\boldsymbol{\theta}}(k-1)\right\} \\ &= \sum_{i=1}^{m}\left\{\tilde{\boldsymbol{\theta}}(k)^T\mathbf{P}_i(k)^{-1}\tilde{\boldsymbol{\theta}}(k) + 2\tilde{\boldsymbol{\theta}}(k-1)^T\mathbf{P}_i(k)^{-1}\sum_{j=1}^{m} k_{mj}(k)\mathbf{P}_j(k)\boldsymbol{\varphi}_j(k)(\boldsymbol{\varphi}_i(k)\tilde{\boldsymbol{\theta}}(k-1) + \sigma_i(k)) - \right. \\ &\quad \left[\sum_{j=1}^{m} k_{mj}(k)\mathbf{P}_j(k)\boldsymbol{\varphi}_j(k)e_j(k)\right]^T \mathbf{P}_i(k)^{-1}\left[\sum_{j=1}^{m} k_{mj}(k)\mathbf{P}_j(k)\boldsymbol{\varphi}_j(k)e_j(k)\right] - \\ &\quad \left.\tilde{\boldsymbol{\theta}}(k-1)^T\mathbf{P}_i(k-1)^{-1}\tilde{\boldsymbol{\theta}}(k-1)\right\}\end{aligned} \tag{22}$$

Considering $\tilde{\boldsymbol{\theta}}(k-1)$ and $\sigma_i(t)$ are orthogonal vectors, let $\mathbf{V}_i(k) = \tilde{\boldsymbol{\theta}}(k)^T\mathbf{P}_i(k)^{-1}\tilde{\boldsymbol{\theta}}(k)$, then (22) can be updated as

$$\begin{aligned}\mathbf{T}(k) = \sum_{i=1}^{m}\left\{\mathbf{V}_i(k) - \mathbf{V}_i(k-1) + 2\tilde{\boldsymbol{\theta}}(k-1)^T\left[\mathbf{P}_i(k)\boldsymbol{\varphi}_i(k)\boldsymbol{\varphi}_i(k)^T k_{mi}(k)\sum_{i=1}^{m}\mathbf{P}_j(k)^{-1}\right]\tilde{\boldsymbol{\theta}}(k-1)\right. \\ \left. - \left[\sum_{i=1}^{m} k_{mi}(k)\mathbf{P}_i(k)\boldsymbol{\varphi}_i(k)e_i(k)\right]^T \mathbf{P}_i(k)^{-1}\left[\sum_{i=1}^{m} k_{mi}(k)\mathbf{P}_i(k)\boldsymbol{\varphi}_i(k)e_i(k)\right]\right\}\end{aligned} \tag{23}$$

Let $\mathbf{W}_i(k) = \tilde{\boldsymbol{\theta}}(k-1)^T\boldsymbol{\varphi}_i(k)\boldsymbol{\varphi}_i(k)^T\tilde{\boldsymbol{\theta}}(k-1)$, then $\mathbf{T}(k)$ can be expressed by $\mathbf{W}_i(k)$

$$\mathbf{T}(k) = \sum_{i=1}^{m}\mathbf{W}_i(k). \tag{24}$$

Take (24) into (23), we have

$$\begin{aligned}\sum_{i=1}^{m}\mathbf{W}_i(k) \le \sum_{i=1}^{m}\left\{\mathbf{V}_i(k) - \mathbf{V}_i(k-1) + 2\mathbf{W}_i(k)\left[\left|\mathbf{P}_i(k)\right|\left|k_{mi}(k)\right|\sum_{j=1}^{m}\left|\mathbf{P}_j(k)^{-1}\right|\right] - \right. \\ \left.\left[\sum_{j=1}^{m} k_{mj}(k)\mathbf{P}_j(k)\boldsymbol{\varphi}_j(k)e_j(k)\right]^T \mathbf{P}_i(t)^{-1}\left[\sum_{j=1}^{m} k_{mj}(k)\mathbf{P}_j(k)\boldsymbol{\varphi}_j(k)e_j(k)\right]\right\}\end{aligned} \tag{25}$$

Similar to the proof of Lemma 2 in Ding and Chen (2004), every item $\mathbf{W}_i(k), i=1,2,...,m$ in (25) can be simplified as

$$\begin{aligned}&\left[\mathbf{I}+\left|\mathbf{P}_i(k)\right|\left|k_{mi}(k)\right|\sum_{j=1}^{m}\left|\mathbf{P}_j(k)^{-1}\right|\right]\mathbf{W}_i(k)\\&\le -2\tilde{\boldsymbol{\theta}}(k)^T\mathbf{P}_i(k-1)^{-1}\sum_{j=1}^{m}k_{mj}(t)\mathbf{P}_j(k-1)\boldsymbol{\varphi}_j(k)\eta_j(k)+\tilde{\boldsymbol{\theta}}(k)^T\boldsymbol{\varphi}_i(k)\boldsymbol{\varphi}_i(k)^T\tilde{\boldsymbol{\theta}}(k)+\\&\left[\sum_{j=1}^{m}k_{mj}(k)\mathbf{P}_j(k)\boldsymbol{\varphi}_j(k)e_j(k)\right]^T\left[\boldsymbol{\varphi}_i(k)^T\boldsymbol{\varphi}_i(k)-2\mathbf{P}_i(k)^{-1}\right]\left[\sum_{j=1}^{m}k_{mj}(k)\mathbf{P}_j(k)\boldsymbol{\varphi}_j(k)e_j(k)\right]\end{aligned}\quad .(26)$$

According the definition of $\mathbf{P}_i(k)$

$$\boldsymbol{\varphi}_i(k)\boldsymbol{\varphi}_i(k)^T-2\mathbf{P}_i(k)^{-1}<0\,. \tag{27}$$

Then the last item in (26) is not positive and take expectation on both sides of (26), we have

$$\begin{aligned}&\left\|\left[1+\left|\mathbf{P}_i(k)\right|\left|k_{mi}(k)\right|\sum_{j=1}^{m}\left|\mathbf{P}_j(k)^{-1}\right|\right]\mathbf{W}_i(k)\right|\le -2\tilde{\boldsymbol{\theta}}(k)^T\mathbf{P}_i(k-1)^{-1}\\&\sum_{i=1}^{m}k_{mi}(k)\mathbf{P}_i(k-1)\boldsymbol{\varphi}_i(k)\eta_i(k)+\left\|\varphi_i(k)\tilde{\boldsymbol{\theta}}(k)\right\|^2\end{aligned}\quad . \tag{28}$$

Considering all inequalities terms about $\mathbf{W}_i(k), i=1,2,...,m$, we have

$$\begin{aligned}&\sum_{i=1}^{m}\left\|\left[1+\left|\mathbf{P}_i(k)\right|\left|k_{mi}(k)\right|\sum_{j=1}^{m}\left|\mathbf{P}_j(k)^{-1}\right|\right]\mathbf{W}_i(k)\right|\\&\le\sum_{i=1}^{m}\left[-2\tilde{\boldsymbol{\theta}}(k)^T\mathbf{P}_i(k-1)^{-1}\sum_{j=1}^{m}k_{mj}(k)\mathbf{P}_j(k-1)\boldsymbol{\varphi}_j(k)\eta_j(k)+\left\|\boldsymbol{\varphi}_i(k)\tilde{\boldsymbol{\theta}}(k)\right\|^2\right]\\&=\sum_{i=1}^{m}\left[-2\tilde{\boldsymbol{\theta}}(k)^T\boldsymbol{\varphi}_i(k)\eta_i(k)\left|k_{mi}(k)\right|\left|\mathbf{P}_i(k-1)\right|\left|\sum_{j=1}^{m}\mathbf{P}_j(k-1)^{-1}\right|+\left\|\boldsymbol{\varphi}_i(k)\tilde{\boldsymbol{\theta}}(k)\right\|^2\right]\end{aligned}\quad .(29)$$

$$\le \sum_{i=1}^{m}\left\{\left\|\boldsymbol{\varphi}_i(k)\tilde{\boldsymbol{\theta}}(k)\right\|^2\left[1-2\left|\mathbf{P}_i(k-1)\right|\sum_{j=1}^{m}\left|\mathbf{P}_j(k-1)^{-1}\right|\left|k_{mi}(k)\right|\right]-\right.$$

$$\left.2\tilde{\boldsymbol{\theta}}(k)^T\boldsymbol{\varphi}_i(k)\left|\mathbf{P}_i(k-1)\right|\sum_{j=1}^{m}\left|\mathbf{P}_j(k-1)^{-1}\right|\left|k_{mi}(k)\right|\sigma_i(k)\right\}$$

$$=\sum_{i=1}^{m}\left\{\left\|\boldsymbol{\varphi}_i(k)\tilde{\boldsymbol{\theta}}(k)\right\|^2\left[1-2\left|\mathbf{P}_i(k-1)\right|\sum_{j=1}^{m}\left|\mathbf{P}_j(k-1)^{-1}\right|\left|k_{mi}(k)\right|\right]-\right.$$

$$2\tilde{\boldsymbol{\theta}}(k-1)\left[1-\boldsymbol{\varphi}_i(k)\sum_{k=1}^{m}k_{mk}(k)\mathbf{P}_k(k)\boldsymbol{\varphi}_k(k)\right]-\sum_{k=1}^{m}k_{mk}(t)\boldsymbol{\varphi}_k(k)^T\mathbf{P}_k(k)\sigma_k(k)$$

$$\left.\left[\mathbf{P}_i(k-1)\sum_{j=1}^{m}\mathbf{P}_j(k-1)^{-1}k_{mi}(k)\boldsymbol{\varphi}_i(k)\sigma_i(k)\right]\right\}$$

As $\sigma_i(k)$ and $\sigma_j(k), i\neq j$ are orthogonal vectors, then (29) can be written as

$$\sum_{i=1}^{m}\left|\left[1+\left|\mathbf{P}_i(k)\right|\left|k_{mi}(k)\right|\sum_{j=1}^{m}\left|\mathbf{P}_j(k)^{-1}\right|\right]\mathbf{W}_i(k)\right|$$

$$\le\sum_{i=1}^{m}\left\|\boldsymbol{\varphi}_i(k)\tilde{\boldsymbol{\theta}}(k)\right\|^2\left[1-2\left|\mathbf{P}_i(k-1)\right|\sum_{j=1}^{m}\left|\mathbf{P}_j(k-1)^{-1}\right|k_{mi}(k)\right]+ \qquad .(30)$$

$$2\sum_{i=1}^{m}\left[\boldsymbol{\varphi}_i(k)^T\mathbf{P}_i(k-1)k_{mi}(k)\sum_{j=1}^{m}k_{mj}(k)\mathbf{P}_j(k)\mathbf{P}_j(k-1)^{-1}\boldsymbol{\varphi}_j(k)\left\|\sigma_j(k)^2\right\|\right]$$

Let $\max\left\|\boldsymbol{\varphi}_i(k)\tilde{\boldsymbol{\theta}}(k)\right\|^2=\bar{u}_{\max}(k), i=1,2,...,m$, and $k_{m1}(k)$ is the minimum among $k_{mi}(k)$, then the item of (30) can be simplified as

$$\sum_{i=1}^{m}\left\|\boldsymbol{\varphi}_i(k)\tilde{\boldsymbol{\theta}}(k)\right\|^2\left[1-2\left|\mathbf{P}_i(k-1)\right|\sum_{j=1}^{m}\left|\mathbf{P}_j(k-1)^{-1}\right|\left|k_{mi}(k)\right|\right]$$

$$\le\bar{u}_{\max}(k)\sum_{i=1}^{m}\left[1-2\left|\mathbf{P}_i(k-1)\right|\sum_{j=1}^{m}\left|\mathbf{P}_j(k-1)^{-1}\right|\left|k_{mi}(k)\right|\right] \qquad .(31)$$

$$\le\bar{u}_{\max}(k)\left[m-2\left|k_{mi}(k)\right|\sum_{i=1}^{m}\sum_{j=1}^{m}\left|\mathbf{P}_i(k-1)\right|\left|\mathbf{P}_j(k-1)^{-1}\right|\right]$$

Because $\mathbf{P}_i(k-1)$ and $\mathbf{P}_j(k-1)$ are full-rank symmetric matrices, we have

$$\begin{cases} \left|\mathbf{P}_i(k-1)\right|\left|\mathbf{P}_j(k-1)^{-1}\right|+\left|\mathbf{P}_j(k-1)\right|\left|\mathbf{P}_i(k-1)^{-1}\right|\geq 2 \\ \mathbf{P}_i(k-1)>0, \mathbf{P}_j(k-1)>0 \end{cases}.$$

then

$$\sum_{i=1}^{m}\left\|\boldsymbol{\varphi}_i(k)\tilde{\boldsymbol{\theta}}(k)\right\|^2\left[1-2\left|\mathbf{P}_i(k-1)\right|\sum_{j=1}^{m}\left|\mathbf{P}_j(k-1)^{-1}\right|k_{mi}(k)\right]\leq \bar{u}_{\max}(k)\left[m-2k_{m1}(k)m^2\right] \quad (32)$$

When $k\to\infty$, $\mathbf{P}_i(k-1)^{-1}\mathbf{P}_i(k)\to\mathbf{I}$, then $\left|k_{m1}(k)\right|\to 1$, as $\bar{u}_{\max}(k)>0$, then the right formula of (32) satisfies

$$\bar{u}_{\max}(k)[m-2k_{m1}(k)m^2]<0. \quad (33)$$

Submitting (32) into (30) and using Lemma1, we have

$$\begin{aligned} &\sum_{i=1}^{m}\left|\left[1+\left|\mathbf{P}_i(k)\right|\left|k_{mi}(k)\right|\sum_{j=1}^{m}\left|\mathbf{P}_j(k)^{-1}\right|\right]\mathbf{W}_i(k)\right| \\ &\leq \sum_{i=1}^{m}2\boldsymbol{\varphi}_i(k)^T\mathbf{P}_i(k-1)\sum_{j=1}^{m}\mathbf{P}_j(k-1)^{-1}\mathbf{P}_j(k)\boldsymbol{\varphi}_j(t)\left\|\sigma_j(k)^2\right\| \\ &<2\sum_{j=1}^{m}\sum_{i=1}^{m}\frac{\boldsymbol{\varphi}_i(k)\boldsymbol{\varphi}_j(k)}{\mathbf{P}_i(k-1)^{-1}}\Delta^2 \\ &<2\sum_{j=1}^{m}\sum_{i=1}^{m}\left(\frac{\beta}{\alpha}+1+\frac{1}{k}\right)\Delta^2 \end{aligned}. \quad (34)$$

Then at time $k\to\infty$, we can get

$$\begin{aligned} \left|\mathbf{T}(k)\right| &< \sum_{i=1}^{m}\left|\left[1+\left|\mathbf{P}_i(k)\right|\left|k_{mi}(k)\right|\sum_{j=1}^{m}\left|\mathbf{P}_j(k)^{-1}\right|\right]\mathbf{W}_i(k)\right| \\ &<2\left(\frac{\beta}{\alpha}+1\right)m^2\Delta^2 \end{aligned}. \quad (35)$$

Appendix C The proof of lemma 3

$$V(k)=\sum_{i=1}^{m}V_i(k)=\sum_{i=1}^{m}\tilde{\boldsymbol{\theta}}(k)^T P_i(k)^{-1}\tilde{\boldsymbol{\theta}}(k)$$

$$=\sum_{i=1}^{m}\left[\tilde{\boldsymbol{\theta}}(k-1)-\sum_{j=1}^{m}k_{mj}(k)\mathbf{P}_j(k)\boldsymbol{\varphi}_j(k)e_j(k)\right]^T\left[\mathbf{P}_i(k-1)^{-1}+\boldsymbol{\varphi}_i(k)\boldsymbol{\varphi}_i(k)^T\right]$$

$$\left[\tilde{\boldsymbol{\theta}}(k-1)-\sum_{j=1}^{m}k_{mj}(k)\mathbf{P}_j(k)\boldsymbol{\varphi}_j(t)e_j(k)\right]$$

$$=\sum_{i=1}^{m}\left\{\tilde{\boldsymbol{\theta}}(k-1)^T\mathbf{P}_i(k-1)^{-1}\tilde{\boldsymbol{\theta}}(k-1)-2\tilde{\boldsymbol{\theta}}(k)^T\mathbf{P}_i(k-1)^{-1}\sum_{j=1}^{m}k_{mj}(k)\mathbf{P}_j(k-1)\boldsymbol{\varphi}_j(k)\eta_j(k)+\right.$$

$$\tilde{\boldsymbol{\theta}}(t)^T\boldsymbol{\varphi}_i(t)\boldsymbol{\varphi}_i(t)^T\tilde{\boldsymbol{\theta}}(t)-\left[\sum_{j=1}^{m}k_{mj}(t)\mathbf{P}_j(t)\boldsymbol{\varphi}_j(k)e_j(k)\right]^T\mathbf{P}_i(k-1)^{-1}$$

$$\left.\left[\sum_{j=1}^{m}k_{mj}(k)\mathbf{P}_j(k)\boldsymbol{\varphi}_j(k)e_j(k)\right]\right\} \tag{36}$$

Considering the definition of $k_{mi}(k)$, then $k_{m1}(k)=\min(k_{mi}(k)), i=1,2,...,m$ and $k_{m1}(k)\in(0,1)$. Moreover, we have

$$\left[\sum_{j=1}^{m}k_{mj}(k)\mathbf{P}_j(k)\boldsymbol{\varphi}_j(k)e_j(k)\right]^T\mathbf{P}_i(k-1)^{-1}\left[\sum_{j=1}^{m}k_{mj}(k)\mathbf{P}_j(k)\boldsymbol{\varphi}_j(k)e_j(k)\right]>0, \tag{37}$$

then

$$V(k)\le V(k-1)-\sum_{i=1}^{m}\left[2\tilde{\boldsymbol{\theta}}(k)^T\mathbf{P}_i(k-1)^{-1}\sum_{j=1}^{m}k_{mj}(k)\mathbf{P}_j(k-1)\boldsymbol{\varphi}_j(k)\eta_j(k)+\tilde{\boldsymbol{\theta}}(k)^T\boldsymbol{\varphi}_i(k)\boldsymbol{\varphi}_i(k)^T\tilde{\boldsymbol{\theta}}(k)\right]$$

$$=V(k-1)-\sum_{i=1}^{m}\left[2\tilde{\boldsymbol{\theta}}(k)^T\mathbf{P}_i(k-1)k_{mi}(k)\boldsymbol{\varphi}_i(k)\eta_i(k)\sum_{j=1}^{m}\mathbf{P}_j(k-1)^{-1}+\tilde{\boldsymbol{\theta}}(k)^T\boldsymbol{\varphi}_i(k)\boldsymbol{\varphi}_i(k)^T\tilde{\boldsymbol{\theta}}(k)\right] \tag{38}$$

Because $\eta_i(k)-\sigma_i(k)$ and $\tilde{\boldsymbol{\theta}}(k)$ are orthogonal, then (38) can be simplified as

$$V(k)\le V(k-1)-|k_{m1}(k)|\left\{\sum_{i=1}^{m}2\tilde{\boldsymbol{\theta}}(k)^T\boldsymbol{\varphi}_i(k)\left(\tilde{\boldsymbol{\theta}}(k)^T\mathbf{P}_i(k-1)\sum_{j=1}^{m}\mathbf{P}_j(k-1)^{-1}\boldsymbol{\varphi}_i(k)+\right.\right.$$

$$\left.\left.\frac{1}{2}\tilde{\boldsymbol{\theta}}(k)^T\boldsymbol{\varphi}_i(k)\right)+2\boldsymbol{\varphi}_i(k)^T\sum_{i=1}^{m}\sum_{j=1}^{m}\mathbf{P}_i(k-1)\mathbf{P}_j(k-1)^{-1}\boldsymbol{\varphi}_i(k)\|\sigma_i(k)\|^2\right\} \tag{39}$$

Let $u_i(k)=\tilde{\boldsymbol{\theta}}(k)^T\boldsymbol{\varphi}_i(k)$, $v_i(k)=\tilde{\boldsymbol{\theta}}(k)^T\mathbf{P}_i(k-1)\sum_{j=1}^{m}\mathbf{P}_j(k-1)^{-1}\boldsymbol{\varphi}_i(k)+\frac{1}{2}\tilde{\boldsymbol{\theta}}(k)^T\boldsymbol{\varphi}_i(k)$, then

$$V(k) \le V(k-1) - 2\left|k_{m1}(k)\right| \sum_{i=1}^{m} u_i(k) v_i(k) + 2m^2 \left\|\sigma(k)\right\|^2 . \tag{40}$$

Obviously, $2\left|k_{m1}(k)\right| \sum_{i=1}^{m} u_i(k) v_i(k) \ge 0$, and set

$$S(k) = 2\sum_{j=1}^{k} \sum_{i=1}^{m} \left|k_{m1}(j)\right| u_i(j) v_i(j) \ge 0 . \tag{41}$$

then

$$\left|V(k) + S(k)\right| \le V(k-1) + S(k-1) + 2m^2 \Delta^2 . \tag{42}$$

Appendix D Explanation for computational complexity in Table 2

Computational complexity is consisted of three parts: data communication, onboard calculation and storage. Because the communications between the head node and leaf nodes are bidirectional with the same communication data amount, we only consider one-way communication from the leaf node to the head. As the computational complexity are determined by matrix inversion $o(N_i^3)$ and multiplication $o(N_i^3)$, we refer Fraleigh and Beauregard, (1995) and Cormen, Leiserson, Rivest, and Stein, (2009) and list the explanations for Table 2 in the following table and the computational complexity of other modes are easily inferred from the calculations of the proposed method:

Table 3: Explanations for Table 2

Tags	Expression	Explanations
Com_ Leaf	KN_p	There are K leaf nodes send N_p identification data to the relay points or to the head node;
Cal_ Leaf	$Ko(N_i^3)$	Leaf nodes calculate the parameter identification by using matrix inversion and matrix multiplication based on the state vector of N_i measuring data;
Cal_ Head	$2N_p$	Head node collects and processes N_p identification data from 2 connected leaf nodes in the circle topology
Sto_Leaf	$K\left(N_i + N_p\right)$	Leaf nodes store N_i measuring data and N_p identification data;
Sto_Head	$2N_p$	Head node stores the identification data N_p from 2 leaf nodes in the circle topology;

* Com = Communication, Cal=Calculation, Sto= Storage;