CYBERNETICS AND PHYSICS, VOL. 14, NO. 2, 2025, 191–199

PARAMETER IDENTIFICATION OF DISCRETE-TIME LINEAR STOCHASTIC SYSTEMS BASED ON DECENTRALIZED SQUARE-ROOT INFORMATION FILTERING

Andrey Tsyganov [1]
[1]Faculty of Mathematics,
Physics and Technology Education
Ulyanovsk State Pedagogical University
Russian Federation
tsyganov.av@ulspu.ru

Julia Tsyganova [2,1]
[2] Institute of Data Science and
Artificial intelligence
Innopolis University
Russian Federation
y.tsyganova@innopolis.ru

Article history:
Received 19.07.2025, Accepted 19.09.2025

Abstract

The paper proposes a new method for identifying parameters of discrete-time linear stochastic systems using decentralized square-root information filtering (DSRIF).

The main contribution of the paper is the derivation of a new identification criterion formulated in terms of DSRIF outputs, such as square roots of information matrices and corresponding estimates of information vectors. An algorithm for its computation is also provided, which uses J-orthogonal transformations at the communication and assimilation stage for updating filter quantities.

The method is validated through a numerical example of circular motion tracking with various configurations of measurement models. Simulations show accurate parameter identification, with improved precision as the number of sensors increases, especially when using sensors measuring the full state vector.

This work establishes a unified framework for decentralized square-root information filtering and parameter identification, suitable for real-life applications in fault-tolerant control, environmental monitoring, and adaptive signal processing.

Key words

discrete-time linear stochastic systems, parameter identification, decentralized estimation, distributed Kalman filter, square-root algorithms, J-orthogonal transformations

1 Introduction

Modern industrial systems increasingly rely on multisensor networks composed of spatially distributed nodes capable of both sensing and local computation. These systems are widely used in diverse applications such as environmental monitoring, air traffic control, and multitarget tracking. The integration of multiple sensors into a communication network enables improved estimation accuracy, robustness, and fault tolerance due to data redundancy and decentralized processing capabilities.

A key architectural advantage of such systems is the ability to implement distributed estimation algorithms, where each node calculates local estimates and shares information with its neighbors in order to achieve a common goal. This approach is particularly effective in large-scale networks where centralized processing is either infeasible or inefficient due to communication constraints and scalability issues.

The Kalman filter remains a cornerstone for state estimation in stochastic systems. However, its direct application in multisensor networks presents several challenges, including high computational demands, communication overhead, and numerical instability due to machine roundoff errors. Moreover, the presence of non-Gaussian noise, model inaccuracies, and unpredictable system behavior further complicates the filtering process. As a result, the development of robust and efficient distributed filtering and parameter estimation algorithms has become a central focus of research in recent years.

In this context, several decentralized filtering approaches have been proposed in the literature. A comparative analysis of such algorithms can be found in [Hidayat et al., 2011], while a comprehensive overview of distributed Kalman filtering techniques is provided in [Mahmoud and Khalid, 2013]. For practical implementation, numerically stable variants of the Kalman filter—such as square-root and information-based formulations—are preferred [Grewal and Andrews, 2015; Tsyganova and Kulikova, 2018]. These implementations are robust to numerical errors and offer advantages in terms of software efficiency and stability. In particular, the square-root filtering technique has been applied in [Tsyganov and Tsyganova, 2019; Radtke et al., 2020; Tsyganov and Tsyganova, 2021] to constructing decentralized discrete-time filtering algorithms.

A promising direction in distributed estimation involves the use of so called consensus-based optimization techniques. For example, the paper [Sergeenko et al., 2021] presents a distributed parameter estimation method using a modified consensus-based SPSA (Simultaneous Perturbation Stochastic Approximation) algorithm. This approach is particularly effective in environments with unknown-but-bounded noise, where traditional statistical assumptions are not applicable.

Similarly, [Erofeeva et al., 2021] introduces a distributed multi-target tracking algorithm that combines SPSA with iterative averaging to address optimization in the presence of signals with uncertain statistical properties. This method is especially suitable for tracking applications where measurement errors and target maneuvers are unpredictable and lack a known probabilistic model. Also, the authors of [Tarasova and Moseiko, 2025] propose a decentralized method based on the SPSA algorithm for adaptive prediction of task processing times in dynamic environments. It enables efficient learning under limited observability and noisy feedback, without requiring access to gradients or global information.

Another important aspect of distributed estimation is the joint estimation of system states and parameters, especially when the system dynamics are not fully known or subject to change. In [Jacobs and DeLaurentis, 2018], a novel methodology is proposed that integrates distributed Kalman filtering with recursive Gaussian process regression for adaptive parameter estimation. This framework allows the network to learn and adapt to unknown or time-varying dynamics, while maintaining consensus among the nodes.

Distributed computations are also used in solving the distributed non-convex resource allocation problem [Fu et al., 2025]. The authors suggest a successive convex approximation-based distributed dual gradient tracking algorithm with a limited communication data rate over a communication network.

Our paper aims to develop a new method for distributed identification of parameters in dynamic systems based on square-root information Kalman filtering. By integrating decentralization mechanisms with robust square-root filtering techniques, we intend to enhance the performance of distributed estimation algorithms in multisensor stochastic systems.

2 Mathematical model of a multisensor measurement data processing system with parameter uncertainty

Consider a mathematical model of a linear discrete-time stochastic system:

$$x_k = F_k x_{k-1} + D_k u_k + G_k w_k \,, \tag{1}$$

$$z_k = H_k x_k + v_k \,, \quad k = 1, 2, \ldots, K \tag{2}$$

where k is a discrete time instant, $x_k \in \mathbb{R}^n$ is the state vector to be estimated; $u_k \in \mathbb{R}^q$ is the control input; $z_k \in \mathbb{R}^m$ is the compound vector of all available measurements at the current time instant k. The noises $\{w_k\}$ and $\{v_k\}$ are independent normally distributed random sequences with zero means and positive definite covariance matrices Q_k and R_k, respectively. The initial value of the state vector is $x_0 \sim \mathcal{N}(\bar{x}_0, \Pi_0)$, which is independent of $\{w_k\}$ and $\{v_k\}$.

Considering the parameter identification problem for the system (1), (2), suppose that all system matrices F_k, D_k, G_k, H_k, covariance matrices Q_k, R_k, Π_0, and also the initial value x_0 can depend on an unknown system model parameter $\theta \in \mathbb{R}^p$ to be identified. It means that $F_k \equiv F_k(\theta)$, $D_k \equiv D_k(\theta)$, etc. But for the sake of simplicity, we will assume, but not specify, the dependence of these quantities on θ.

Further, we assume that the measurement model (2) is multisensory. The latter means that the measurement matrix H_k and the error covariance matrix R_k can be represented as

$$H_k = [H_{1,k}^T | H_{2,k}^T | \ldots | H_{N,k}^T]^T, \tag{3}$$

and

$$R_k = \text{blockdiag}(R_{1,k},\ R_{2,k},\ \ldots,\ R_{N,k}). \tag{4}$$

Thus, equation (2), taking into account (3) and (4), is a global multisensory measurement model which also depends on θ.

Consider a fully connected network of sensors consisting of N nodes, in which each node i has the ability to compute the estimates $\hat{x}_{i,k}$ of the state vector x_k and the corresponding error covariance matrices $P_{i,k}$. The measurements and estimates obtained at each node are called local [Rao and Durrant-Whyte, 1991].

Suppose that the state equation (1) is the same in each node, and local measurements are described by the following equation

$$z_{i,k} = H_{i,k} x_k + v_{i,k} \tag{5}$$

where $v_{i,k} \sim \mathcal{N}(0, R_{i,k})$, $i = 1, \ldots, N$. Equation (4) implies that the measurement noises at nodes i and j are uncorrelated.

Since the state vector x_k is not directly observable, discrete-time Kalman filtering algorithms are usually used to estimate it. By now, there exist many different modifications of the Kalman filter [Grewal and Andrews, 2015; Tsyganova and Kulikova, 2018]. There are classes of so called covariance algorithms in which at each step the error covariance matrix P_k is updated (the conventional Kalman filter belongs to this class). Also, there are classes of information algorithms in which, instead of P_k and x_k, the values of the information matrix $Y_k \triangleq P_k^{-1}$ and the information state vector estimate $y_k \triangleq Y_k x_k$ are computed at each step. Information algorithms, due to their structure, are most preferred for decentralized processing of measurement data [Rao and Durrant-Whyte, 1991]. Consider the information Kalman filtering algorithm, which consists of two stages [Grewal and Andrews, 2015]:

Algorithm 1. Information Kalman filter (IKF).
For $k = 1, 2, \ldots, K$
I. Time update

$$A_k = F_k^{-T}\hat{Y}_{k-1}F_k^{-1}, \quad (6)$$
$$C_k = G_k^T A_k G_k + Q_k^{-1}. \quad (7)$$
$$L_k = A_k G_k C_k^{-1}, \quad (8)$$
$$\tilde{y}_k = [I - L_k G_k^T]F_k^{-T} \times \left(\hat{y}_{k-1} + \hat{Y}_{k-1}F_k^{-1}D_k u_k\right), \quad (9)$$
$$\tilde{Y}_k = [I - L_k G_k^T]A_k, \quad (10)$$

II. Measurement update

$$\hat{y}_k = \tilde{y}_k + \Delta y_k,\ \Delta y_k = H_k^T R_k^{-1} z_k, \quad (11)$$
$$\hat{Y}_k = \tilde{Y}_k + \Delta Y_k,\ \Delta Y_k = H_k^T R_k^{-1} H_k \quad (12)$$

where Δy_k and ΔY_k are updates of the information vector $\hat{y}_k$ and the information matrix $\hat{Y}_k$, correspondingly.
End For

Equations (6)–(10), together with (11), (12), are derived from the conventional Kalman filter taking into account the definition of the information matrix and information state estimates, as well as application of matrix inversion lemmas.

3 Decentralized measurement data processing

The key idea of decentralized measurement data processing is the ability to express global updates of the information vector and the information matrix through local [Hidayat et al., 2011; Rao and Durrant-Whyte, 1991]:

$$\Delta y_k = H_k^T R_k^{-1} z_k = \sum_{i=1}^{N} H_{i,k}^T R_{i,k}^{-1} z_{i,k} = \sum_{i=1}^{N} \Delta y_{i,k}, \quad (13)$$

$$\Delta Y_k = H_k^T R_k^{-1} H_k = \sum_{i=1}^{N} H_{i,k}^T R_{i,k}^{-1} H_{i,k} = \sum_{i=1}^{N} \Delta Y_{i,k}. \quad (14)$$

Local updates are computed at each node and transmitted to all other nodes.

Thus, the decentralized information filter contains three following stages.

Algorithm 2. Decentralized information filter (DIF).
For $k = 1, 2, \ldots, K$
I. Local time update

$$\tilde{y}_{i,k} = [I - L_k G_k^T]F_k^{-T} \times \left(\hat{y}_{i,k-1} + \hat{Y}_{i,k-1}F_k^{-1}D_k u_k\right), \quad (15)$$
$$\tilde{Y}_{i,k} = [I - L_k G_k^T]A_k \quad (16)$$

where A_k and L_k are computed according to (6)–(8).
II. Local measurement update

$$\Delta y_{i,k} = H_{i,k}^T R_{i,k}^{-1} z_{i,k}, \quad (17)$$
$$\Delta Y_{i,k} = H_{i,k}^T R_{i,k}^{-1} H_{i,k}^T. \quad (18)$$

III. Communication and assimilation

$$\hat{Y}_{i,k} = \tilde{Y}_{i,k} + \sum_{j=1}^{N} \Delta Y_{j,k}, \quad (19)$$
$$\hat{y}_{i,k} = \tilde{y}_{i,k} + \sum_{j=1}^{N} \Delta y_{j,k}. \quad (20)$$

Here $i = 1, \ldots, N$. At any discrete time instant k, each estimate $\hat{x}_{i,k} = \hat{Y}_{i,k}^{-1}\hat{y}_{i,k}$ of the state vector x_k is available at every node of the multisensor network consisting of N nodes.
End For

As shown in [Rao and Durrant-Whyte, 1991], a decentralized filter (15)–(20) is algebraically equivalent to the centralized (conventional) Kalman filter with the measurement model (2).

4 Decentralized square-root information filtering

The Square-Root Information Filter (SRIF) was first proposed in [Dyer and McReynolds, 1969]. The authors used matrix square roots, that is, the representation of the information matrix Y in the form $Y = S^T S$ where S is the Cholesky factor (an upper triangular matrix). The vector $s = Sx$ is called the square root information state vector [Grewal and Andrews, 2015, p. 357]. The SRIF algorithm is widely used due to its improved computational properties compared to the standard implementation of the information filter.

A recent paper [Tsyganov and Tsyganova, 2021] proposed a new method for decentralized processing of measurement data using a square-root information filter.

Suppose that at the stages of local time and measurement update in the square-root information algorithm local square-root information estimates $\tilde{s}_{i,k}$, $\hat{s}_{i,k}$ and the matrix square roots $\tilde{S}_{i,k}$, $\hat{S}_{i,k}$ are computed. Then stages I (equations (15), (16)) and II (equations (17), (18)) in the decentralized information filter for zero control input can be written as follows ($i = 1, \ldots, N$).

Algorithm 3. Decentralized square-root information filter (DSRIF).
For $k = 1, 2, \ldots, K$
I. Local time update

$$\tilde{\mathcal{Q}}_i \begin{bmatrix} Q_k^{-1/2} & 0 & 0 \\ -\hat{S}_{i,k-1}F_k^{-1}G_k & \hat{S}_{i,k-1}F_k^{-1} & \hat{s}_{i,k-1} \end{bmatrix} = \begin{bmatrix} (*) & (*) & (*) \\ 0 & \tilde{S}_{i,k} & \tilde{s}_{i,k} \end{bmatrix}$$

where $\tilde{\mathcal{Q}}_i$ is the matrix of orthogonal transformation to upper triangular form for the block matrix on the left-hand side of this equation, $Q_k^{-1/2}$ is the Cholesky square root of covariance matrix Q_k^{-1}, here and after $(*)$ denotes matrix blocks that are not of interest.

In the case of non-zero control input u_t, it is necessary to adjust the estimate $\tilde{s}_{i,k}$ as follows:

$$\tilde{s}_{i,k} := \tilde{s}_{i,k} + \hat{S}_{i,k-1}D_k u_k.$$

II. Local measurement update

$$\hat{\mathcal{Q}}_i \begin{bmatrix} \tilde{S}_{i,k} & \tilde{s}_{i,k} \\ R_{i,k}^{-1/2}H_{i,k} & R_{i,k}^{-1/2}z_{i,k} \end{bmatrix} = \begin{bmatrix} \hat{S}_{i,k} & \hat{s}_{i,k} \\ 0 & e_{i,k} \end{bmatrix}, \quad (21)$$

where $\hat{\mathcal{Q}}_i$ is the matrix of orthogonal transformation to the upper triangular form of the block matrix on the left-hand side of this equation, $R_{i,k}^{-1/2}$ is the Cholesky square root of covariance matrix $R_{i,k}^{-1}$, and $e_{i,k}$ is the residual of measurement $z_{i,k}$.
End For

When constructing a decentralized square-root information filter, the main difficulty arises during the formulation of the communication and assimilation stage. In order to implement this, the authors of [Roy et al., 1991] proposed using matrix orthogonal transformations in complex-valued arithmetic. However, this approach can significantly complicate the software implementation of the algorithm and slow down computational speed.

Unlike the approach supposed in [Roy et al., 1991], we suggest using a J-orthogonal transformation of the form $\bar{\mathcal{Q}}A = R$ for the effective implementation of communication and assimilation stages. Here the transformation matrix $\bar{\mathcal{Q}}$ is J-orthogonal, that is, $\bar{\mathcal{Q}}^T J \bar{\mathcal{Q}} = J$, J is a signature matrix of the form $J = (I_p \oplus -I_q)$ ($p \geq 1$, $q \geq 1$) [Higham, 2003].

This idea was first proposed in [Tsyganov and Tsyganova, 2019], which presents the formulation of the communication and assimilation stage in the decentralized square root information filtering algorithm, and later, in [Tsyganov and Tsyganova, 2021], a rigorous theoretical justification for the developed algorithm was given.

From (20) it follows that $\Delta Y_{i,k} = \hat{Y}_{i,k} - \tilde{Y}_{i,k}$ and $\Delta y_{i,k} = \hat{y}_{i,k} - \tilde{y}_{i,k}$, $i = 1, \ldots, N$. Since in the SRIF $Y = S^T S$, $s = Sx$, and at the same time $y = S^T S x = S^T s$, then $\Delta Y_{i,k} = \hat{S}_{i,k}^T \hat{S}_{i,k} - \tilde{S}_{i,k}^T \tilde{S}_{i,k}$ and $\Delta y_{i,k} = \hat{S}_{i,k}^T \hat{s}_{i,k} - \tilde{S}_{i,k}^T \tilde{s}_{i,k}$.

Let us write down the communication and assimilation stage (20) in terms of the SRIF (let $\langle \cdot \rangle_i$ be the values computed at the i-th iteration).

III. Communication and assimilation
A. Set $\langle \hat{S}_k \rangle_0 = \tilde{S}_k$, $\langle \hat{s}_k \rangle_0 = \tilde{s}_k$.
B. For $i = 1, 2, \ldots, N$ do

$$\bar{\mathcal{Q}}_i \begin{bmatrix} \langle \hat{S}_k \rangle_{i-1} & \langle \hat{s}_k \rangle_{i-1} \\ \hat{S}_{i,k} & \hat{s}_{i,k} \\ \tilde{S}_{i,k} & \tilde{s}_{i,k} \end{bmatrix} = \begin{bmatrix} \langle \hat{S}_k \rangle_i & \langle \hat{s}_k \rangle_i \\ 0 & (*) \\ 0 & (*) \end{bmatrix} \quad (22)$$

where $\tilde{S}_{i,k}$, $\hat{S}_{i,k}$ and $\tilde{s}_{i,k}$, $\hat{s}_{i,k}$ are the Cholesky square roots of the information matrices and the information square root state vector estimates, obtained at the i-th node at the stages of local time and measurement update using the square-root information filtering algorithm; $\bar{\mathcal{Q}}_i$ is the matrix of J-orthogonal transformation from the first column of the block matrix on the left-hand side of (22) to the upper triangular form, with $\bar{\mathcal{Q}}_i^T J \bar{\mathcal{Q}}_i = J$ where $J = (I_p \oplus -I_q)$ ($p = 2n$, $q = n$).

C. Obtain the result $\hat{S}_k = \langle \hat{S}_k \rangle_N$, $\hat{s}_k = \langle \hat{s}_k \rangle_N$.

A detailed proof of stage III can be found in [Tsyganov and Tsyganova, 2021].

5 Construction of a new identification criterion based on the decentralized square-root information filter

Since we are considering a class of discrete-time linear stochastic systems with additive Gaussian noises, we propose to construct a new parameter identification criterion based on a negative log-likelihood function. This function has the following structure:

$$J(\theta, Z_1^K) = \frac{Km}{2}\ln(2\pi) + \frac{1}{2}\sum_{k=1}^{K}\left[\ln\det B_k + \|\nu_k\|^2_{B_k^{-1}}\right] \quad (23)$$

where the vector of residuals ν_k and its covariance matrix B_k are calculated using the conventional Kalman algorithm [Gibbs, 2011].

Our main goal is to develop a new identification method based on a decentralized square-root information filter. In order to achieve this, we need to be able to calculate the value of $\ln\det B_k$ and the squared norm of ν_k, using the output data from the DSRIF algorithm.

Let us formulate the main result obtained.

Theorem 1. *The identification criterion (23) in terms of a decentralized square-root information filter DSRIF has the form*

$$\begin{aligned} J_{DSRIF}(\theta, Z_1^K) &= \frac{Km}{2}\ln(2\pi) \\ &+ \frac{1}{2}\sum_{k=1}^{K}\Bigg[2\left(\sum_{i=1}^{N}\ln\det R_{i,k}^{1/2} + \ln\det \tilde{S}_k - \ln\det \hat{S}_k\right) \\ &\qquad + \|\tilde{s}_k\|^2 - \|\hat{s}_k\|^2 + \sum_{i=1}^{N}\|R_{i,k}^{-1/2} z_{i,k}\|^2\Bigg] \end{aligned} \quad (24)$$

where the right-hand side of (24) *contains the values available in the DSRIF at each step of its execution* ($k = 1, \ldots, K$).

Proof. We need to prove that

$$\ln\det B_k = 2\left(\sum_{i=1}^{N}\ln\det R_{i,k}^{1/2} + \ln\det\tilde{S}_k - \ln\det\hat{S}_k\right) \quad (25)$$

and

$$\|\nu_k\|^2_{B_k^{-1}} = \|\tilde{s}_k\|^2 - \|\hat{s}_k\|^2 + \sum_{i=1}^{N}\|R_{i,k}^{-1/2} z_{i,k}\|^2. \quad (26)$$

Let us use one of the results from the theory of square-root discrete filtering to consider the measurement update stage of a square-root covariance filter [Tsyganova and Kulikova, 2018, p. 11]:

$$\mathcal{Q}\begin{bmatrix} S_{R_k} & H_k S_{P_k} \\ 0 & S_{P_k}\end{bmatrix} = \begin{bmatrix} S_{R_{e,k}} & 0 \\ \bar{K}_{f,k} & S_{P_{k|k}}\end{bmatrix}. \quad (27)$$

Here S_{R_k} is the Cholesky square root of the measurement errors covariance matrix R_k, matrices $S_{P_{k|k}}$, S_{P_k} are the Cholesky square roots of the estimation error covariance matrices at the stages of time update and measurement update, respectively; $S_{R_{e,k}}$ is the Cholesky square root of the measurement residual covariance matrix, $\bar{K}_{f,k} = P_k H_k^T S_{R_{e,k}}^{-1}$.

Since (27) uses the Cholesky decomposition in the form $P = S_P S_P^T$, we can write the following relations between matrices, taking into account the notation introduced:

$$\begin{aligned} &S_{R_k} \triangleq R_k^{T/2},\ S_{P_{k|k}} \triangleq \tilde{S}_k^{-T}, \\ &S_{P_k} \triangleq \hat{S}_k^{-T},\ S_{R_{e,k}} \triangleq S_{B_k}^T. \end{aligned} \quad (28)$$

Let us rewrite (27) taking into account (28):

$$\mathcal{Q}\begin{bmatrix} R_k^{T/2} & H_k\hat{S}_k^{-T} \\ 0 & \hat{S}_k^{-T}\end{bmatrix} = \begin{bmatrix} S_{B_k}^T & 0 \\ \bar{K}_{f,k} & \tilde{S}_k^{-T}\end{bmatrix}. \quad (29)$$

Considering that $\mathcal{Q}$ is an orthogonal matrix, we can write the correct equality

$$\begin{aligned} &\begin{bmatrix} R_k^{1/2} & 0 \\ \hat{S}_k^{-1}H_k^T & \hat{S}_k^{-1}\end{bmatrix}\cdot \mathcal{Q}^T\mathcal{Q}\cdot\begin{bmatrix} R_k^{T/2} & H_k\hat{S}_k^{-T} \\ 0 & \hat{S}_k^{-T}\end{bmatrix} \\ &\qquad = \begin{bmatrix} S_{B_k} & \bar{K}_{f,k}^T \\ 0 & \tilde{S}_k^{-1}\end{bmatrix}\cdot\begin{bmatrix} S_{B_k}^T & 0 \\ \bar{K}_{f,k} & \tilde{S}_k^{-T}\end{bmatrix}. \end{aligned} \quad (30)$$

Taking into account the properties of the determinant of orthogonal and block triangular matrices, from (30) we obtain

$$\begin{aligned} &\det R_k^{1/2}\det\hat{S}_k^{-1}\det R_k^{T/2}\det\hat{S}_k^{-T} \\ &\qquad = \det S_{B_k}\det\tilde{S}_k^{-1}\det S_{B_k}^T\det\tilde{S}_k^{-T}. \end{aligned} \quad (31)$$

Thus,

$$\det B_k = \left(\frac{\det R_k^{1/2}\det\hat{S}_k^{-1}}{\det\tilde{S}_k^{-1}}\right)^2. \quad (32)$$

Considering that for any square non-singular matrix $\det A^{-1} = \frac{1}{\det A}$, equation (32) can be rewritten as

$$\det B_k = \left(\frac{\det R_k^{1/2}\det\tilde{S}_k}{\det\hat{S}_k}\right)^2. \quad (33)$$

Taking into account (33) and the well-known properties of logarithms, we arrive at

$$\ln\det B_k = 2\left(\ln\det R_k^{1/2} + \ln\det\tilde{S}_k - \ln\det\hat{S}_k\right). \quad (34)$$

Since matrix R_k has a block diagonal structure (4), we can represent $\ln\det R_k^{1/2}$ as

$$\ln\det R_k^{1/2} = \sum_{i=1}^{N}\ln\det R_{i,k}^{1/2}. \quad (35)$$

Substituting (35) into (34), we obtain (25). Thus, expression (25) is proved.

Now let us prove (26). Consider expression (21) in general form:

$$\hat{\mathcal{Q}}\begin{bmatrix} \tilde{S}_k & \tilde{s}_k \\ R_k^{-1/2}H_k & R_k^{-1/2}z_k\end{bmatrix} = \begin{bmatrix} \hat{S}_k & \hat{s}_k \\ 0 & e_k\end{bmatrix}, \quad (36)$$

Considering that $\hat{\mathcal{Q}}$ is an orthogonal matrix and $\hat{\mathcal{Q}}^T\hat{\mathcal{Q}} = I$, we can write the correct equality

$$\begin{aligned} &\begin{bmatrix} \tilde{S}_k^T & H_k^T R_k^{-T/2} \\ \tilde{s}_k^T & z_k^T R_k^{-T/2}\end{bmatrix}\begin{bmatrix} \tilde{S}_k & \tilde{s}_k \\ R_k^{-1/2}H_k & R_k^{-1/2}z_k\end{bmatrix} \\ &\qquad = \begin{bmatrix} \hat{S}_k^T & 0 \\ \hat{s}_k^T & e_k^T\end{bmatrix}\begin{bmatrix} \hat{S}_k & \hat{s}_k \\ 0 & e_k\end{bmatrix}, \end{aligned} \quad (37)$$

from where

$$\tilde{s}_k^T \tilde{s}_k + z_k^T R_k^{-T/2} R_k^{-1/2} z_k = \hat{s}_k^T \hat{s}_k + e_k^T e_k. \quad (38)$$

The resulting equality (38) can be written as

$$\|e_k\|^2 = \|\tilde{s}_k\|^2 - \|\hat{s}_k\|^2 + \|R_k^{-1/2} z_k\|^2. \quad (39)$$

Next, we use the previously proven fact [Bierman et al., 1990] that $\|\nu_k\|^2_{B_k^{-1}} = \|e_k\|^2$.

Thus,

$$\|\nu_k\|^2_{B_k^{-1}} = \|\tilde{s}_k\|^2 - \|\hat{s}_k\|^2 + \|R_k^{-1/2} z_k\|^2. \quad (40)$$

Taking into account the block diagonal structure (4) of the matrix R_k,

$$\|R_k^{-1/2} z_k\|^2 = \sum_{i=1}^{N} \|R_{i,k}^{-1/2} z_{i,k}\|^2. \quad (41)$$

Substituting (41) into (40), we arrive at (26).

So, the theorem 1 has been fully proved.

The constructed criterion (24) can be used to solve the problem of parameter identification for the system (1), (2) in the case where the input signal u_k is known or equal zero. In the case of an unknown input signal, the criterion proposed in [Tsyganova and Tsyganov, 2023] can be used to solve the parameter identification problem.

6 Algorithm for calculating the new identification criterion based on a decentralized square-root information filter

In this section, we propose an algorithm for calculating the identification criterion (24). To do this, we modify the DSRIF algorithm, extending its functionality with the ability to calculate the values of the identification criterion $J_{DSRIF}(\theta, Z_1^K, i)$ in each ith node of the multisensor network.

First, let us denote

$$J_{i,k}^{(1)} = \sum_{i=1}^{N} \left(\ln\det R_{i,k}^{1/2} + \frac{1}{2}\|z_{i,k}\|^2_{R_{i,k}^{-1}} \right), \quad (42)$$

$$J_{i,k}^{(2)} = \ln\det \tilde{S}_k + \frac{1}{2}\|\tilde{s}_k\|^2, \quad (43)$$

$$J_{i,k}^{(3)} = \ln\det \hat{S}_k + \frac{1}{2}\|\hat{s}_k\|^2, \quad (44)$$

and then rewrite (24) in a more convenient form

$$J_{DSRIF}(\theta, Z_1^K) = \frac{Km}{2}\ln(2\pi) + \sum_{k=1}^{K}\left[J_{i,k}^{(1)} + J_{i,k}^{(2)} - J_{i,k}^{(3)} \right]. \quad (45)$$

The result is presented by algorithm 4.

Algorithm 4. Calculating the value of the identification criterion $J_{DSRIF}(\theta, Z_1^K, i)$ in the ith node

Initialization

$J_{DSRIF}(\theta, Z_1^K, i) = \frac{Km}{2}\ln(2\pi)$

For $k = 1, 2, \ldots, K$

I. Local time update

$$\tilde{\mathcal{Q}}_i \begin{bmatrix} Q_k^{-1/2} & 0 & 0 \\ -\hat{S}_{i,k-1} F_k^{-1} G_k & \hat{S}_{i,k-1} F_k^{-1} & \hat{s}_{i,k-1} \end{bmatrix} = \begin{bmatrix} (*) & (*) & (*) \\ 0 & \tilde{S}_{i,k} & \tilde{s}_{i,k} \end{bmatrix}$$

$J_{i,k}^{(2)} = \ln\det \tilde{S}_{i,k} + \frac{1}{2}\|\tilde{s}_{i,k}\|^2$

II. Local measurement update

$$\hat{\mathcal{Q}}_i \begin{bmatrix} \tilde{S}_{i,k} & \tilde{s}_{i,k} \\ R_{i,k}^{-1/2} H_{i,k} & R_{i,k}^{-1/2} z_{i,k} \end{bmatrix} = \begin{bmatrix} \hat{S}_{i,k} & \hat{s}_{i,k} \\ 0 & e_{i,k} \end{bmatrix}$$

III. Communication and assimilation

$\langle \hat{S}_k \rangle_0 = \tilde{S}_{i,k}$, $\langle \hat{s}_k \rangle_0 = \tilde{s}_{i,k}$, $J_{i,k}^{(1)} = 0$

For $j = 1, 2, \ldots, N$

$$\bar{\mathcal{Q}}_i \begin{bmatrix} \langle \hat{S}_k \rangle_{j-1} & \langle \hat{s}_k \rangle_{j-1} \\ \hat{S}_{j,k} & \hat{s}_{j,k} \\ \tilde{S}_{j,k} & \tilde{s}_{j,k} \end{bmatrix} = \begin{bmatrix} \langle \hat{S}_k \rangle_j & \langle \hat{s}_k \rangle_j \\ 0 & (*) \\ 0 & (*) \end{bmatrix}$$

$J_{i,k}^{(1)} := J_{i,k}^{(1)} + \ln\det R_{j,k}^{1/2} + \frac{1}{2}\|z_{j,k}\|^2_{R_{j,k}^{-1}}$

End For

$\hat{S}_{i,k} = \langle \hat{S}_k \rangle_N$, $\hat{s}_{i,k} = \langle \hat{s}_k \rangle_N$.

$J_{i,k}^{(3)} = \ln\det \hat{S}_{i,k} + \frac{1}{2}\|\hat{s}_{i,k}\|^2$

$J_{DSRIF}(\theta, Z_1^K, i) := J_{DSRIF}(\theta, Z_1^K, i) + J_{i,k}^{(1)} + J_{i,k}^{(2)} - J_{i,k}^{(3)}$

End For

At stage III of algorithm 4, $\tilde{S}_{j,k}$ and $\hat{S}_{jk}$ are square roots of information matrices; vectors $\tilde{s}_{j,k}$ and $\hat{s}_{j,k}$ are corresponding state estimates obtained in the jth node at stages of local time update and local measurement update; $\bar{\mathcal{Q}}_i$ is the matrix of J-orthogonal transformation that brings the first column of block matrix on the left side of (22) into upper triangular form.

7 Numerical example

Consider a model of circular clockwise motion of an object moving on a plane, described by the following equation:

$$x_k = F x_{k-1} + D u_k + G w_k \quad (46)$$

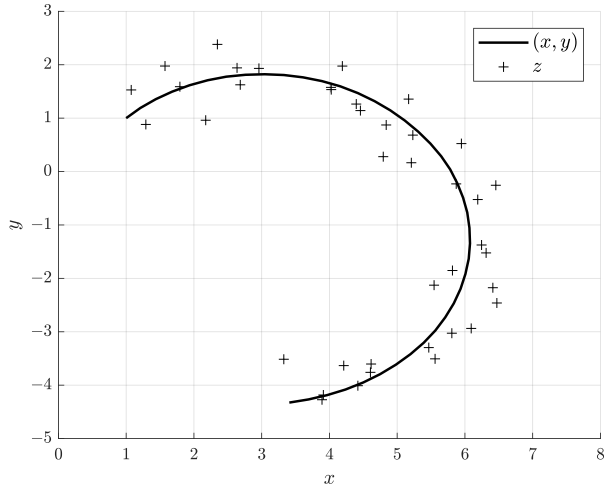

Figure 1. Trajectory and noisy measurements.

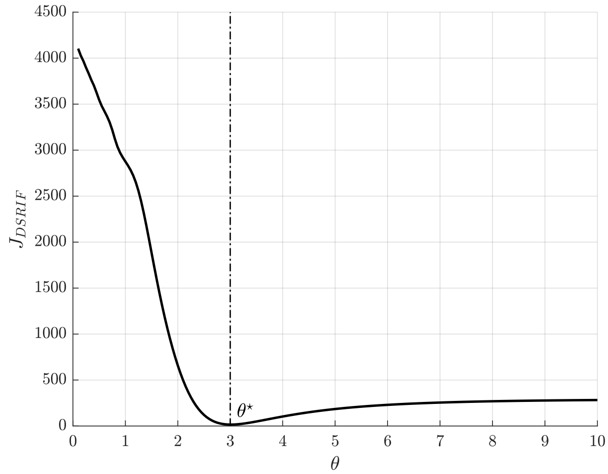

Figure 2. Identification criterion.

Table 1. Identification results.

Sensors	Mean	RMSE	MAPE
$[1]$	3.0039	0.0496	1.3055
$[1, 1]$	3.0005	0.0428	1.1430
$[2]$	3.0013	0.0463	1.2307
$[2, 2]$	3.0008	0.0380	0.9985
$[3]$	2.9989	0.0370	0.9905
$[3, 3]$	3.0014	0.0296	0.8046
$[1, 2, 3]$	3.0002	0.0323	0.8565

where

$$F = \begin{bmatrix} \Phi & 0 \\ 0 & \Phi \end{bmatrix}, \; \Phi = \begin{bmatrix} \cos\omega\tau & \omega^{-1}\sin\omega\tau \\ -\omega\sin\omega\tau & \cos\omega\tau \end{bmatrix},$$

$$D = \begin{bmatrix} (x_{1,0} + \omega^{-1}x_{4,0})(1-\cos\omega\tau) \\ (\omega x_{1,0} + x_{4,0})\sin\omega\tau \\ (x_{3,0} - \omega^{-1}x_{2,0})(1-\cos\omega\tau) \\ (\omega x_{3,0} - x_{2,0})\sin\omega\tau \end{bmatrix}, \; G = \begin{bmatrix} 0 & 0 \\ 1 & 0 \\ 0 & 0 \\ 0 & 1 \end{bmatrix},$$

$x_k = \left[x_1, \, x_2, \, x_3, \, x_4\right]_k^T$ is the state vector, τ is the sampling interval, $\omega = |v_0|/r > 0$, r is the given radius, $v_0 = \begin{bmatrix} x_{2,0} \\ x_{4,0} \end{bmatrix}$ is the velocity vector at the initial point with coordinates $(x_{1,0}, x_{3,0})$, $u_{k-1} \equiv 1$, $w_k \sim \mathcal{N}(0, Q)$ (see, for example, [Semushin et al., 2017]).

Let $\tau = 0.1$, $x_0 = \left[1, \, 2, \, 1, \, 2\right]^T$, $Q = 0.001 I_2$, $\theta = r$ and the true value of the parameter $\theta^\star = 3$. Consider the following sensors:

$$z_{i,k} = H_{i,k} x_k + v_{i,k}, \quad i = 1, 2, 3 \tag{47}$$

where

$$H_{1,k} = \begin{bmatrix} 1 & 0 & 0 & 0 \end{bmatrix}, \quad R_{1,k} = 0.1;$$

$$H_{2,k} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix}, \quad R_{2,k} = \begin{bmatrix} 0.1 & 0 \\ 0 & 0.1 \end{bmatrix};$$

$$H_{3,k} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}, \quad R_{3,k} = \begin{bmatrix} 0.1 & 0 & 0 & 0 \\ 0 & 0.1 & 0 & 0 \\ 0 & 0 & 0.1 & 0 \\ 0 & 0 & 0 & 0.1 \end{bmatrix}.$$

The modeling of the object's motion and the identification process was performed in MATLAB. Figure 1 shows the graphs of the object's trajectory and noisy measurements, and Figure 2 shows the graph of the identification criterion for the second sensor and $K = 40$.

The minimization of the identification criterion was performed using the standard function `fmincon`, for which the objective function of calculating the identification criterion was implemented, as well as a number of auxiliary functions. The search for the estimate of the parameter θ was performed on the interval $[0.1, 10]$ using the results of 40 measurements, with the middle of the interval selected as the initial approximation. Table 1 shows the mean value of the found estimates of the parameter θ, the root mean square error (RMSE) and the mean absolute percentage error (MAPE) for different sets of sensors based on the results of 500 experiments.

The results of numerical experiments show that the unknown parameter is identified correctly both for measurement models with one sensor and for measurement models with several sensors. For multisensor measurement models, the results are identical across all network nodes. When the number of sensors increases, the accuracy of unknown parameter identification also increases. The measurement models that include sensors for measuring all components of the state vector demonstrate high accuracy of identification results.

8 Conclusion

The paper proposes a new method for constructing parameter identification algorithms based on decentralized square-root information filtering. The method is presented in the form of Algorithm 4 which has the ability to calculate the values of an identification criterion simultaneously at each node in a multisensor network. Numerical optimization methods are also used to minimize the criterion (24) and find estimates of the unknown parameter θ. The principal theoretical result is Theorem 1 on a new identification criterion in terms of a decentralized square-root information filtering.

The novelty of the proposed method is that a decentralized square-root information algorithm for distributed Kalman filtering has been chosen for its implementation. The difference between this algorithm and previously proposed distributed or parallel square-root computing schemes is that the stage of communication and assimilation is based on the J-orthogonal transformation of a block data array. As a result, the square root of the information vector estimate and the corresponding square root of the information matrix are calculated using the values received from each node in the multisensor network.

The proposed algorithm has the ability for decentralized calculation of an identification criterion. It can be used in decentralized computational schemes for estimating a state vector of a dynamic system based on multisensor measurement data, under conditions of prior uncertainty about the parameters of a mathematical model of the system.

It should be noted that to ensure the correct minimization of the identification criterion, the key point of the presented algorithm is a consistent exploration of the search space by all nodes of the network, which requires a fully connected network to obtain identical values of the identification criterion in each node. Adaptation of the proposed approach to partially connected networks is the subject of future research.

Acknowledgements

This work was carried out under the Supplementary Agreement No. 073-03-2025-066/1 dated March 19, 2025, to the Agreement on granting subsidies from the federal budget for financial support of the fulfillment of state assignments to provide public services (works) No. 073-03-2025-066 dated January 16, 2025, concluded between Ulyanovsk State Pedagogical University and the Ministry of Education of the Russian Federation.

References

Bierman, G. J., Belzer, M. R., Vandergraft, J. S., and Porter, D. W. (1990). Maximum likelihood estimation using square root information filters. *IEEE Transactions on Automatic Control*, **35** (12), pp. 1293–1298.

Dyer, P. and McReynolds, S. (1969). Extension of square-root filtering to include process noise. *J. Optimiz. Theory Appl.*, (3), pp. 444–459.

Erofeeva, V., Granichin, O., Granichina, O., Proskurnikov, A., and Sergeenko, A. (2021). Weighted SPSA-based consensus algorithm for distributed cooperative target tracking. In *2021 European Control Conference (ECC)*, pp. 1074–1079.

Fu, K., Fu, X., and Li, T. (2025). Distributed dual gradient tracking for nonconvex resource allocation with limited communication data rate. *Cybernetics and Physics*, **14** (1), pp. 40–51.

Gibbs, B. P. (2011). *Advanced Kalman filtering, least-squares and modeling: a practical handbook*. John Wiley & Sons, Inc., Hoboken, New Jersey.

Grewal, M. S. and Andrews, A. P. (2015). *Kalman filtering: Theory and Practice Using MATLAB*. John Wiley & Sons, Inc., 4th edition.

Hidayat, Z., Babuška, R., De Schutter, B., and Núñez, A. (2011). Decentralized Kalman filter comparison for distributed-parameter systems: A case study for a 1D heat conduction process. In *ETFA2011*, pp. 1–8.

Higham, N. J. (2003). J-orthogonal matrices: Properties and generation. *SIAM Review*, **45** (3), pp. 504–519.

Jacobs, M. A. and DeLaurentis, D. (2018). Distributed Kalman filter with a Gaussian process for machine learning. In *2018 IEEE Aerospace Conference*, pp. 1–12.

Mahmoud, M. and Khalid, H. (2013). Distributed Kalman filtering: A bibliographic review. *IET Control Theory and Applications*, **7** (4), pp. 483–501.

Radtke, S., Noack, B., and Hanebeck, U. D. (2020). Fully decentralized estimation using square-root decompositions. In *2020 IEEE 23rd International Conference on Information Fusion (FUSION)*, pp. 1–8.

Rao, B. S. and Durrant-Whyte, H. F. (1991). Fully decentralised algorithm for multisensor Kalman filtering. *IEE Proceedings D — Control Theory and Applications*, **138** (5), pp. 413–420.

Roy, S., Hashemi, R. H., and Laub, A. J. (1991). Square root parallel Kalman filtering using reduced-order local filters. *IEEE Transactions on Aerospace Electronic Systems*, **27** (2), pp. 276–289.

Semushin, I. V., Tsyganov, A. V., Tsyganova, Y. V., Golubkov, A. V., and Vinokurov, S. D. (2017). Modelling and estimation of a moving object trajectory. *Vestnik YuUrGU. Ser. Mat. Model. Progr.*, **10** (3), pp. 108–119. (in Russian).

Sergeenko, A., Erofeeva, V., Granichin, O., Granichina, O., and Proskurnikov, A. (2021). Convergence analysis of weighted SPSA-based consensus algorithm in distributed parameter estimation problem. *IFAC-PapersOnLine*, **54** (7), pp. 126–131. 19th IFAC Symposium on System Identification SYSID 2021.

Tarasova, E. and Moseiko, E. (2025). Decentralized SPSA-based correction of task time predictions in adaptive multi-agent systems. *Cybernetics and Physics*, **14** (1), pp. 67–73.

Tsyganov, A. V. and Tsyganova, Y. V. (2019). Decentralized square-root information algorithm of multisensor Kalman filtering. In *Proc. of the XIII All-Russian*

Meeting on Control Problems VSPU-2019, Moscow, Russia, June 17–20, pp. 679–683. (in Russian).

Tsyganov, A. V. and Tsyganova, Y. V. (2021). Decentralized square-root information Kalman filtering based on J-orthogonal transformations. *Journal of Physics: Conference Series*, **1745** (1), pp. 012112.

Tsyganova, Y. and Tsyganov, A. (2023). Parameter identification of the linear discrete-time stochastic systems with unknown exogenous inputs. *Cybernetics and Physics*, **12** (3), pp. 219–229.

Tsyganova, Y. V. and Kulikova, M. V. (2018). On modern array algorithms for optimal discrete filtering. *Vestnik YuUrGU. Ser. Mat. Model. Progr.*, **11** (4), pp. 5–30. (in Russian).