This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication.
Citation information: DOI 10.1109/TAC.2025.3589276, IEEE Transactions on Automatic Control

Remote State Estimation for Nonlinear Systems Under Compression-Decompression Mechanism: A Modified Unscented Kalman Filtering Approach

Jiahao Song, Zidong Wang, Qinyuan Liu, and Xiao He

***Abstract*—In engineering practice, some large-scale systems have high-dimensional measurements that exhibit redundancy and are suitable to be compressed. Measurement compression-decompression is an effective approach to saving communication resources in networked control systems, and compressive sensing (CS) is a popular high-performance compression-decompression method for such measurements. In this paper, we investigate the remote state estimation for nonlinear systems under a compression-decompression mechanism on the measurement output. With the application of CS, a state estimator is designed based on the unscented Kalman filter. Despite the prominent advantages of CS, the presence of measurement noise and quantization errors in practice is inevitable, which could lead to a degradation in the performance of CS and an enlargement of state estimation errors. To address this challenge, we analyze the combined influence of measurement noise and quantization errors on the performance of data compression-decompression and state estimation. The design of estimator gains is approached by minimizing an upper bound of the estimation error covariance. Furthermore, a sufficient condition is derived to ensure the mean-square exponential boundedness of the estimation error. Finally, the effectiveness of the proposed method is verified through simulation experiments conducted on power grid systems, which are characterized by highly redundant measurements that are suitable for compression-decompression.**

***Index Terms*—State estimation, networked system, unscented Kalman filter, data compression, compressive sensing.**

I. Introduction

In the past few decades, networked control systems (NCSs) have received ever-growing attention from both academia and industry owing primarily to the rapid advancement of network technologies. The introduction of networks into control systems has brought several remarkable advantages including flexible architecture, simplified deployment, and reduced maintenance costs [16], [31], [42]. These superiorities have enabled the extensive applications of NCSs in various scenarios such as power grids, environment monitoring, and industrial automation [7], [22], [30], [33]. During the implementation and operation of NCSs, state estimation has been recognized as an essential task whose aim is to obtain precise estimates of system state information. State estimation is crucial for several key tasks like control, decision-making, fault diagnosis, and so on [44], [46], [49]. Consequently, state estimation for NCSs has been identified as an important research domain in recent years [12], [34], [45].

The limitation of communication resources has been identified as one of the most important challenges in the application of NCSs [19], [40], [41], restricting the amount of information that can be transmitted and thus degrading the performance of state estimation. To handle this challenge, compression-decompression of measurement data has been a widely applied technique in recent years [47], [48]. Specifically, the fundamental task of data compression involves the reduction of data size while retaining essential information, whereas data decompression aims at recovering the original information from the compressed data [14], [21]. In this paper, a particular interest is shown in large-scale practical systems with high-dimensional state vectors and measurement vectors with strong correlations between the vectors' components. Typical examples can be seen in power grids, transportation systems, antenna arrays, and so on [2], [4], [10]. Such high dimensions and strong correlations indicate extensive redundancy in measurement vectors, which are suitable for compression-decompression.

Compressive sensing (CS), a promising compression-decompression technique attracting extensive attention over the past two decades, is particularly effective for processing data characterized by high dimensions and high redundancy. Several remarkable advantages are associated with CS, including a high compression rate, efficient data acquisition, and low computational complexity for compression, among others [25]. As a result, CS has found widespread application in various scenarios such as sonar systems, radar systems, geophysical sensing systems, image compression, data mining, Magnetic Resonance Imaging, and others [17], [25]. Given the significant strengths of CS demonstrated in engineering practice, the employment of CS in state estimation tasks to achieve high-performance data compression-decompression

This work was supported in part by National Natural Science Foundation of China under grants 62473223, 62222312, and 62473285, in part by Beijing Natural Science Foundation under grant L241016, in part by the National Key Research and Development Program of China under Grant 2022YFB4501704, in part by the Shanghai Science and Technology Innovation Action Plan Project of China under Grant 22511100700, in part by Fundamental Research Funds for the Central Universities, in part by the China Scholarship Council under grant 202206210302, in part by the Royal Society of the U.K., and in part by the Alexander von Humboldt Foundation of Germany. *(Corresponding author: Xiao He)*

Jiahao Song and Xiao He are with the Department of Automation, Tsinghua University, Beijing 100084, China. (Emails: songjh20@mails.tsinghua.edu.cn;hexiao@tsinghua.edu.cn).

Zidong Wang is with the Department of Computer Science, Brunel University London, Uxbridge UB8 3PH, United Kingdom. (Email: Zidong.Wang@brunel.ac.uk).

Qinyuan Liu is with the Department of Computer Science and Technology, Tongji University, Shanghai 201804, China, the Key Laboratory of Embedded System and Service Computing, Ministry of Education, Tongji University, Shanghai 200092, China, and the Shanghai Artificial Intelligence Laboratory, Shanghai, China. (Email: liuqy@tongji.edu.cn).

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication.
Citation information: DOI 10.1109/TAC.2025.3589276, IEEE Transactions on Automatic Control

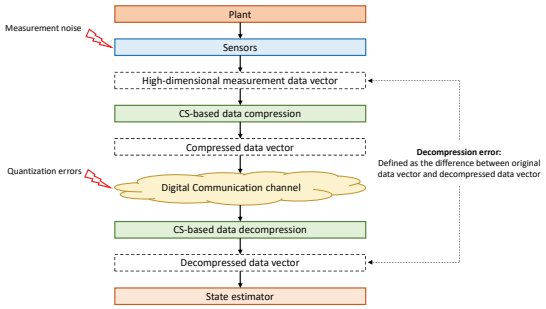


Fig. 1: A fundamental workflow of CS-based state estimation

has been a prevalent research topic [15], [23]. Therefore, the main focus of this paper is on incorporating CS into state estimation. A fundamental workflow of CS-based state estimation is illustrated in Fig. 1, where the core idea involves employing CS to convey measurement data while conserving communication resources.

Despite the excellent performance and promising prospects of CS, it is important to acknowledge that CS is a lossy compression-decompression method, that is, the data obtained after decompression is not identical to the original data prior to compression. Here, decompression error is defined as the difference between the decompressed data and the original data. Such decompression error is unavoidable and is influenced by various factors. A notable issue in state estimation is the inevitability of measurement noise. Measurement noise compromises the level of data redundancy, thereby rendering the measurement data more challenging to compress and decompress. Given that CS is particularly sensitive to such redundancy degradation, measurement noise can lead to large decompression errors and the loss of key information, which is detrimental to state estimation. Thus, it becomes a major challenge to analyze the influence of measurement noise on data compression-decompression in CS-based state estimation.

Another notable issue arises from the necessity of data quantization due to the prevalent use of digital communication channels in NCSs. As illustrated in Fig. 1, compressed data must undergo quantization before transmission through the communication channel. Data quantization inevitably introduces quantization errors into the transmitted data. Compared to uncompressed data, compressed data possesses lower dimensions and less redundancy, and is therefore more susceptible to the effects of quantization errors. Therefore, the negative impact of quantization errors must be considered when implementing data compression-decompression methods. Unfortunately, analyzing the influence of data quantization on CS presents a particularly challenging task as the decompression process of CS typically involves complex iterative optimization steps [6], [23], [29].

In recent years, there has been a surge in the literature on assisting state estimation with CS-based data compression. For instance, when system states possess high dimensions and embedded redundancy, it has been shown in [3] that the measurement vectors can be transformed into sparse vectors, where the sparsity can characterize the redundancy, and are thus particularly suitable to be compressed by CS. Based on this fundamental example, some well-established state estimation approaches (i.e. Kalman filtering and Bayesian filtering) have been integrated with CS [11], [15], [23]. A general case, where special measurement models lead to measurement vectors with high redundancy, has also been investigated by implementing CS for measurement compression [28]. However, to the best of our knowledge, a theoretical analysis of the impact of measurement noise and quantization effects on the performance of CS-based state estimation remains largely unexplored.

In light of the discussions presented above, the key challenges confronted include: 1) the application of CS for the compression and decompression of noisy measurement data, followed by the development of a state estimator based on the decompressed measurements; 2) the analysis of the combined impact of measurement noise and quantization errors on the data compression-decompression process; and 3) the design of state estimator gains to ensure the performance of state estimation under the influence of data compression-decompression. With the aim to address these challenges, this paper focuses primarily on the state estimation problem with measurement compression-decompression by taking into consideration unfavorable factors such as measurement noise and quantization errors. The contributions of this paper are highlighted as follows.

1) The data decompression error is characterized by analyzing the influence of measurement noise and quantization errors based on a well-established CS approach for data compression-decompression.
2) A modified unscented Kalman filter (UKF) is integrated with measurement compression-decompression and implemented for state estimation, wherein the filter gains are designed to minimize an upper bound of the estimation error covariance.
3) The impact of data compression-decompression on state estimation performance is analyzed, and a sufficient condition is derived for ensuring the mean-square exponential boundedness of the estimation error.

II. Problem Formulation and Preliminaries

In Fig. 2, a block diagram is shown to illustrate the architecture of the state estimation system considered in this paper. In this system, the compression-decompression process is implemented using CS, digital communication is taken into account, and state estimation is performed utilizing data after decompression. In this section, details will be provided to model this state estimation system.

A. Redundancy in Measurement Data

As discussed in the previous section, the aim here is to conserve communication resources by implementing data compression-decompression through utilizing the redundancy inherent in the measurement data. It is crucial to articulate the nature of this redundancy and ascertain the suitability of the measurement data for compression-decompression.

First, the concept of sparsity is defined as follows.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication.
Citation information: DOI 10.1109/TAC.2025.3589276, IEEE Transactions on Automatic Control

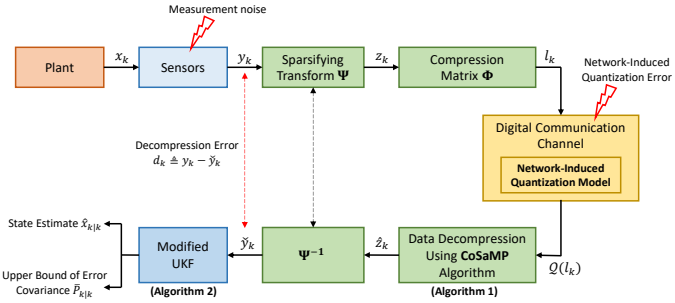


Fig. 2: Block diagram of the state estimation system

Definition 1 (s-sparsity): A vector is called s-sparse if it has no more than s non-zero components.

In a sparse vector, the quantity of non-zero components is typically significantly lower than the vector's dimension. The bulk of essential information within a sparse vector is conveyed by its non-zero components, creating a distinct division between useful and redundant information in sparse vectors. Consequently, the concept of sparsity is particularly apt for characterizing the redundancy present in measurement vectors.

To uncover redundancy and highlight key information in measurement vectors, it is imperative to establish a linkage between the redundancy in measurement data and sparsity. In certain practical situations, vectors characterized by redundancy can be transformed into sparse vectors through appropriate transformations [28]. This transformation process is termed sparsification, which is formally defined as follows.

Definition 2 (s-sparsification): For a vector x and a matrix Ψ, if Ψx is an s-sparse vector, we say that Ψ can s-sparsify x.

In accordance with the discussion above, a measurement vector amenable to sparsification is deemed suitable for compression and decompression. Subsequently, the data that has been decompressed can be employed for state estimation. This paper is dedicated to providing a rigorous mathematical description of both the compression-decompression process and the ensuing state estimation process. Furthermore, and most critically, the impact of compression-decompression on the performance of state estimation will be subjected to further analysis.

B. System Model

Consider a class of discrete-time nonlinear dynamical systems as follows:

$$x_{k+1} = f(x_k, u_k) + w_k, \tag{1}$$

where $x_k \in \mathbb{R}^n$ is the state variable, $f(\cdot) : \mathbb{R}^n \to \mathbb{R}^n$ is a known nonlinear state transition function, u_k represents the known input vector, and $w_k \in \mathbb{R}^n$ denotes Gaussian white noise with zero mean and known covariance Q_k.

In this paper, a special type of system is considered wherein the measurement vectors exhibit considerable redundancy and are amenable to sparsification. An illustrative example of such a system is the smart grid system as investigated in [28]. To fulfill the condition for sparsification, the measurement model must adhere to certain complex conditions. For the sake of concise expression, such systems are characterized by a general nonlinear measurement model as follows:

$$y_k = h(x_k) + v_k, \tag{2}$$

where $y_k \in \mathbb{R}^m$ is the measurement vector, $h(\cdot) : \mathbb{R}^n \to \mathbb{R}^m$ is a known nonlinear measurement function, and $v_k \in \mathbb{R}^m$ represents Gaussian white noise with zero mean and known covariance R_k.

To aid in unveiling the redundancy inherent in the measurement, the following assumption is established.

Assumption 1: There exist a non-singular matrix $\Psi \in \mathbb{R}^{m\times m}$ and a constant $s \in \mathbb{N}^+$ $(s \ll m)$ such that the noise-free measurement, namely $h(x_k)$, can always be s-sparsified by Ψ.

A matrix Ψ, which satisfies Assumption 1, plays a pivotal role in revealing the redundancy inherent in measurement vectors. Herein, such a matrix Ψ is defined as a *sparsifying transformation*.

It should be emphasized that the choice of Ψ is not unique. As mentioned in [28], for certain practical systems, various transformations (e.g. Fourier transform and wavelet transforms) can potentially act as sparsifying transformations [27]. To verify whether Assumption 1 is satisfied, one can try out different sparsifying transformations to see if a suitable candidate for Ψ can be found. In practice, Assumption 1 is usually satisfied for systems with redundant or correlated measurements. In this paper, it is assumed that a suitable sparsifying transformation Ψ is already known.

C. CS-Based Measurement Compression

Denote

$$z_k^* \triangleq \Psi(y_k - v_k) = \Psi h(x_k), \tag{3}$$

$$z_k \triangleq \Psi y_k = z_k^* + \Psi v_k. \tag{4}$$

In accordance with Definition 2 and Assumption 1, z_k^* is classified as an s-sparse vector. However, due to the unknown nature of the process noise v_k, the true value of z_k^* remains unattainable. Given the availability of z_k, the objective is to perform data compression on z_k. Note that z_k is not necessarily s-sparse due to the impact of the process noise v_k. We will conduct a detailed analysis of the influence of the process noise later.

In this paper, given that CS is a widely recognized high-performance technique for the compression and decompression of sparse vectors, we choose to employ CS to process z_k. The fundamental principle of CS involves employing a linear mapping to transform a high-dimensional sparse vector into a low-dimensional non-sparse vector. By applying CS, the compressed data vector, represented by l_k, can be expressed as

$$l_k = \Phi z_k, \tag{5}$$

where $\Phi \in \mathbb{R}^{p\times m}$ is a predefined matrix and $p \ll m$ is a known constant integer.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication.
Citation information: DOI 10.1109/TAC.2025.3589276, IEEE Transactions on Automatic Control

Remark 1: As the compressed low-dimensional vector l_k is unable to convey as much information as the high-dimensional vector z_k, it is imperative that the matrix Φ is selected with care to ensure that the compression process does not result in excessive loss of critical information. The criteria for selecting Φ will be elaborated later in Lemma 1.

D. Network-Induced Quantization

In the context of the widely utilized digital communication channels, data quantization is a prerequisite prior to data transmission. Furthermore, since quantization is implemented to the compressed data, the influence of quantization is non-negligible to the state estimation process. Thus, quantization is a necessary step in the data transmission process.

In this part, an element-wise uniform quantization, which is commonly applied, is considered for the compressed data vector l_k. Denoting $l_k = \begin{bmatrix} l_{k,1} & l_{k,2} & \cdots & l_{k,p} \end{bmatrix}^T$, the quantizer is described as:

$$\mathcal{Q}(l_k) \triangleq \begin{bmatrix} \mathcal{Q}_1(l_{k,1}) & \mathcal{Q}_2(l_{k,2}) & \cdots & \mathcal{Q}_p(l_{k,p}) \end{bmatrix}^T.$$

Here, $\mathcal{Q}_i(\cdot)$ $(i \in \{1, 2, \cdots, p\})$ are uniform quantizers for scalars and $\mathcal{Q}_i(\cdot)$ is given as

$$\mathcal{Q}_i(l_{k,i}) \triangleq \delta_i \mathcal{R}\left(\frac{l_{k,i}}{\delta_i}\right),$$

where $\delta_i > 0$ is a pre-defined quantization level and $\mathcal{R}(\cdot)$ stands for the function rounding a real number to the nearest integer. As a result, the quantization error of each element satisfies

$$|l_{k,i} - \mathcal{Q}_i(l_{k,i})| \leqslant \frac{\delta_i}{2},\ i = 1, 2, \cdots, p. \tag{6}$$

Denoting the quantization error as

$$\varepsilon_k \triangleq \mathcal{Q}(l_k) - l_k, \tag{7}$$

the result of quantization can be rewritten as

$$\mathcal{Q}(l_k) = \Phi z_k + \varepsilon_k.$$

Denote

$$\bar{\varepsilon} \triangleq \frac{1}{2}\sqrt{\sum_{i=1}^{p} \delta_i^2}.$$

Using (6), it is clear that the quantization error vector satisfies

$$\|\varepsilon_k\|_2 \leqslant \bar{\varepsilon}.$$

Though not being the main focus of this paper, limited quantization levels may also be a practical concern. In this case, we can consider clipping the measurement vector y_k so that the compressed data vector l_k does not exceed the range of quantization. The impact of clipping is the same as that of sensor saturation, which has already been extensively studied in the existing literature [36].

E. CS-Based Measurement Decompression

In this paper, it is assumed that the quantized data $\mathcal{Q}(l_k)$ can be transmitted through the digital communication channel without any loss. Consequently, data decompression becomes necessary. Adhering to the principles of CS, the decompression of measurements can be undertaken by solving the following optimization problem [6], [13]:

$$\begin{aligned} \hat{z}_k^* \triangleq\ & \underset{\hat{z}_k}{\operatorname{argmin}} \|\mathcal{Q}(l_k) - \Phi \hat{z}_k\|_2 \\ & \text{s.t. } \|\hat{z}_k\|_0 \leqslant s, \end{aligned} \tag{8}$$

where $\|\hat{z}_k\|_0$ stands for the l_0 norm of $\hat{z}_k$, which represents the count of non-zero components within the vector $\hat{z}_k$. The main idea of this optimization problem is to derive an s-sparse vector $\hat{z}_k^*$ that is capable of reconstructing the received data vector $\mathcal{Q}(l_k)$ with maximal accuracy.

Due to the non-convex nature of the l_0 norm, the optimization problem (8) is classified as NP-hard [9]. As a feasible alternative, the Compressive Sampling Matching Pursuit (CoSaMP) algorithm is employed to secure an approximated solution [29]. In this context, CoSaMP$_s$ is denoted as the CoSaMP algorithm specifically for obtaining an s-sparse vector.

Define $\operatorname{supp}(x)$ as the support of a vector x, which represents a set containing the indices of non-zero elements of x. For a set T, $|T|$ represents the number of elements in T. For $z \in \mathbb{R}^m$ and $T \subset \{1, 2, \cdots, n\}$, denote $z|_T \in \mathbb{R}^{|T|}$ as the part of elements in z with indices in T. Similarly, for a matrix $\Phi \in \mathbb{R}^{p \times m}$, $\Phi|_T \in \mathbb{R}^{p \times |T|}$ is composed of the columns of Φ with indices in T. Furthermore, denote z_s as an s-sparse vector obtained from the vector z by preserving the s elements with the largest magnitudes and setting all other elements to zero. Similarly, y_{2s} can also be obtained by setting all elements in y to zero except the $2s$ elements with the largest magnitudes. b_s can be acquired in the same way. Moreover, denote $A^\dagger$ as the Moore-Penrose pseudo inverse of the matrix A. With the above notations, the CoSaMP$_s$ algorithm for measurement decompression can be summarized as Algorithm 1.

Algorithm 1 CoSaMP$_s$ for measurement decompression [29]

Input: Compressed and quantized data vector $\mathcal{Q}(l_k)$, compression matrix Φ, number of iterations t_k, and allowed number of non-zero components s.

For initialization, let $\hat{z}_k^{[0]} \leftarrow \mathbf{0}$ and $v \leftarrow \mathcal{Q}(l_k)$.
for $k \leftarrow 1$ to t_k **do**
  $y \leftarrow \Phi^T v$.
  $\Omega \leftarrow \operatorname{supp}(y_{2s})$.
  $T \leftarrow \Omega \cup \operatorname{supp}\left(\hat{z}^{[k-1]}\right)$.
  $b|_T \leftarrow \Phi_T^\dagger \mathcal{Q}(l_k)$.
  $b|_{T^C} \leftarrow \mathbf{0}$.
  $\hat{z}^{[k]} \leftarrow b_s$.
  $v \leftarrow \mathcal{Q}(l_k) - \Phi \hat{z}_{[k]}$.
end for
Let $\hat{z}_k \leftarrow \hat{z}_k^{[t_k]}$.

Output: $\hat{z}_k$ as an approximated solution to (8).

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication.
Citation information: DOI 10.1109/TAC.2025.3589276, IEEE Transactions on Automatic Control

Using $\mathcal{Q}(l_k)$, CoSaMP$_s$ tries to recover z_k and produces $\hat{z}_k$. In light of (4), the decompressed measurement, denoted by $\check{y}_k$, can be acquired with $\check{y}_k = \Psi^{-1}\hat{z}_k$.

Unfortunately, the result of decompression $\check{y}_k$ is different from the original measurement vector y_k due to the following three reasons: 1) CS is a lossy compression-decompression method; 2) CoSAMP$_s$ only acquires an approximated solution of (8); and 3) measurement noise and quantization errors pose additional disturbances on the compression-decompression process. Here, we denote the difference between $\check{y}_k$ and y_k as the decompression error:

$$d_k \triangleq y_k - \check{y}_k. \tag{9}$$

Apart from $\mathcal{Q}(l_k)$ and Φ, CoSaMP$_s$ is reliant on two other important parameters, namely the number of non-zero elements s and the number of iterations t_k. Here, s is a known constant that is related to the characteristics of the system model. t_k influences the decompression error, and will be decided later based on requirements on the decompression performance.

The performance of the CoSaMP algorithm is dependent on the dimensions and sparsity of z_k. Generally, a more sparse vector z_k with higher dimensions leads to a smaller decompression error. Though a precise relation is currently unavailable, as a rough guideline, CoSaMP would have a good performance if the order of magnitude of p satisfies

$$p \geqslant Cs\log\frac{m}{s}$$

where C indicates an algorithm-related constant. Therefore, CS-based compression-decompression is typically suitable for high-dimensional systems in which the measurement vectors exhibit redundancy and sparsity.

Remark 2: An intuitive method to deal with data compression is to directly incorporate the compression process into the measurement model, which involves considering

$$z_k = \Psi\left(h\left(x_k\right) + v_k\right) \tag{10}$$

as a modified measurement equation. Based on (10), a state estimator can then be designed. This method appears to be easily implementable and can circumvent the complex decompression operation. However, this approach is flawed because it neglects the critical aspect of sparsity, thereby compromising observability. Given that a nonlinear system is under consideration in this paper, a theoretical analysis of observability is unfeasible. Furthermore, the disadvantage of using (10) as the measurement model for state estimation will be demonstrated through simulation experiments in section IV.

F. State Estimation Problem

It is worth noting that the remote estimator can only receive the decompressed measurement $\check{y}_k$, which is inevitably distorted by noises. Thus, the state estimation problem requires us to utilize $\check{y}_k$, the system model (1)–(2), information about the compression-decompression process, and knowledge of the noises to design a state estimator that reaches optimal estimation performance.

To help specify the performance index for state estimation, we first construct the estimator. To handle the nonlinear system characteristics in (1)–(2), a UKF with the following update rule is implemented following [24]:

$$\hat{x}_{k+1|k+1} = \hat{x}_{k+1|k} + K_{k+1}\left(\check{y}_{k+1} - \hat{y}_{k+1|k}\right), \tag{11}$$

where the notations are slightly different from the filter design in [24]. Here, $\hat{x}_{k+1|k}$ is the one-step prediction at time instant $k+1$, $\hat{x}_{k+1|k+1}$ is the state estimate at time instant $k+1$, $\check{y}_{k+1}$ stands for the decompressed measurement, $\hat{y}_{k+1|k}$ represents the predicted measurement based on $\hat{x}_{k+1|k}$, and K_{k+1} denotes the filter gain matrix. $\hat{x}_{k+1|k}$ and $\hat{y}_{k+1|k}$ can be obtained by following the standard UKF method, while K_{k+1} is to be designed later.

To help design and analyze the state estimator, the state prediction error, state estimation error, and measurement prediction error at time instant $k+1$ are denoted, respectively, as follows:

$$\tilde{x}_{k+1|k} \triangleq x_{k+1} - \hat{x}_{k+1|k} \tag{12}$$

$$\tilde{x}_{k+1|k+1} \triangleq x_{k+1} - \hat{x}_{k+1|k+1} \tag{13}$$

$$\tilde{y}_{k+1|k} \triangleq y_{k+1} - \hat{y}_{k+1|k}. \tag{14}$$

Based on the above notations, we further denote

$$P_{k+1|k} \triangleq \mathbb{E}\left\{\tilde{x}_{k+1|k}\tilde{x}_{k+1|k}^T\right\}$$
$$P_{k+1|k+1} \triangleq \mathbb{E}\left\{\tilde{x}_{k+1|k+1}\tilde{x}_{k+1|k+1}^T\right\}$$
$$P_{xy,k+1|k} \triangleq \mathbb{E}\left\{\tilde{x}_{k+1|k}\tilde{y}_{k+1|k}^T\right\}$$
$$P_{yy,k+1|k} \triangleq \mathbb{E}\left\{\tilde{y}_{k+1|k}\tilde{y}_{k+1|k}^T\right\},$$

where $P_{k+1|k}$ is the covariance of the state prediction error, $P_{k+1|k+1}$ denotes the covariance of the state estimation error, $P_{xy,k+1|k}$ stands for the cross-covariance of state prediction error and measurement prediction error, and $P_{yy,k+1|k}$ represents the covariance of measurement prediction error. In this paper, $\mathrm{tr}\left\{P_{k+1|k+1}\right\}$ is selected as the main performance index characterizing the overall estimation accuracy.

As will be clarified later, it is infeasible to analytically compute the exact expression of $P_{k+1|k+1}$ owing to the combined impact of quantization errors and measurement compression-decompression. As a tractable alternative, an upper bound of $P_{k+1|k+1}$, denoted by $\bar{P}_{k+1|k+1}$, is to be obtained. Subsequently, the state estimator gain matrix K_{k+1} should be designed to minimize $\mathrm{tr}\left\{\bar{P}_{k+1|k+1}\right\}$ to help improve the estimation precision. Finally, the performance of the designed estimator should be analyzed by evaluating the mean-square exponential boundedness of the estimation error.

The main results, which are presented in the next section, are in correspondence with the problems and objectives stated above. Specifically, the impact of the compression-decompression process on the data transmission process is analyzed in detail in Section III-A. In Section III-B, a comprehensive analysis is proposed to illustrate how the data transmission process influences the estimation error, and $\bar{P}_{k+1|k+1}$ is subsequently obtained. Filter gains are also designed to minimize $\mathrm{tr}\left\{\bar{P}_{k+1|k+1}\right\}$. At last, performance analysis is conducted

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication.
Citation information: DOI 10.1109/TAC.2025.3589276, IEEE Transactions on Automatic Control

for the designed estimator, where conditions are proposed for ensuring the mean-square exponential boundedness of the estimation error in Section III-C.

III. Main Results

A. Analysis of Decompression Error

In this part, we aim to analyze the combined influence of measurement noise and quantization errors on the decompression error.

We start by studying a basic decompression problem. Consider a compressed vector $y = \Phi x + e$, where x is the original data vector, Φ is a matrix with fewer rows than columns, and e denotes bounded disturbance. Denote x_s as an s-sparse vector obtained by preserving s components with the largest magnitudes in x and setting all other components to zero. Here, x_s is a best s-sparse approximation to x [29], which satisfies

$$\begin{aligned} x_s = &\underset{\bar{x}}{\operatorname{argmin}} \left\|x - \bar{x}\right\|_2 \\ &\text{s.t. } \left\|\bar{x}\right\|_0 \leqslant s. \end{aligned} \tag{15}$$

Let

$$\tilde{\epsilon}_s \triangleq \left\|x - x_s\right\|_2 + \frac{1}{\sqrt{s}} \left\|x - x_s\right\|_1 + \left\|e\right\|_2. \tag{16}$$

Let $x^{[t]}$ represent the outcome obtained after executing the CoSaMP$_s$ for t iterations. Then, the efficacy of CoSAMP$_s$ can be described by the following lemma.

Lemma 1 (Theorem 2.1, [29]): If there exists a constant $0 < \delta_{4s} \leqslant 0.1$ such that the inequality

$$(1 - \delta_{4s}) \left\|z\right\|_2^2 \leqslant \left\|\Phi z\right\|_2^2 \leqslant (1 + \delta_{4s}) \left\|z\right\|_2^2 \tag{17}$$

holds for any vector z satisfying $\left\|z\right\|_0 \leqslant 4s$, then there exists a positive constant κ so that

$$\left\|x - x^{[t]}\right\|_2 \leqslant 2^{-t} \left\|x_s\right\|_2 + \kappa \tilde{\epsilon}_s. \tag{18}$$

Remark 3: The constant κ in (18) is dependent on a few factors such as the matrix Φ and the specific value of δ_{4s} in (17). Variations in system characteristics result in differing values of κ. The readers are referred to [29] for further elaboration. Using random matrices as candidates is an effective method for choosing Φ [8].

It is pointed out in Lemma 1 that the upper bound of $\left\|x - x^{[t]}\right\|_2$ decreases and converges to $\kappa\tilde{\epsilon}_s$ as the number of iterations t increases. Specifically speaking, the term $2^{-t}\left\|x_s\right\|_2$ in (18) converges to zero while t approaches infinity. Therefore, we can always keep the term $2^{-t}\left\|x_s\right\|_2$ below a predefined level $\theta > 0$ by selecting a proper t. With θ given as a known constant, Lemma 1 can be extended as follows:

Lemma 2: Consider the decompression process introduced in Lemma 1. Suppose that the number of iterations t in Lemma 1 is set as

$$t^* \triangleq \lceil \log_2 \frac{\left\|x\right\|_2}{\theta} \rceil \tag{19}$$

where, for any $a \in \mathbb{R}$, $\lceil a \rceil$ indicates the smallest integer equal to or larger than a. Then, the decompression error satisfies

$$\left\|x - x^{[t^*]}\right\|_2 \leqslant \theta + \kappa\tilde{\epsilon}_s$$

with $\tilde{\epsilon}_s$ defined in (16).

Proof: Combining (18) and (19) yields

$$\begin{aligned} \left\|x - x^{[t^*]}\right\|_2 \leqslant &2^{-t^*} \left\|x_s\right\|_2 + \kappa\tilde{\epsilon}_s \\ \leqslant &\frac{\theta}{\left\|x\right\|_2} \left\|x\right\|_2 + \kappa\tilde{\epsilon}_s \\ = &\theta + \kappa\tilde{\epsilon}_s, \end{aligned}$$

which completes the proof. ∎

According to Lemma 2, when obtaining $\hat{z}_k$ using Algorithm 1, we can set the number of iterations t_k as

$$t_k = \lceil \log_2 \frac{\left\|z_k\right\|_2}{\theta} \rceil. \tag{20}$$

While z_k remains unknown during the decompression phase, it is accessible during the compression process, enabling the straightforward calculation of t_k. Subsequently, t_k can be relayed through the communication network together with the compressed data vector. Given that t_k is merely an integer, its transmission imposes minimal strain on the communication network.

With t_k decided, we can employ CoSaMP$_s$ for data decompression. It is quite challenging to analyze the characteristics of $z_k - \hat{z}_k$ or $y_k - \check{y}_k$ due mainly to the existence of the measurement noise and quantization errors. To address this issue, we present Lemma 3 and Theorem 1.

Lemma 3: x_s satisfying (15) is also a solution to the following optimization problem:

$$\begin{aligned} &\underset{\bar{x}}{\operatorname{argmin}} \left\|x - \bar{x}\right\|_1 \\ &\text{s.t. } \left\|\bar{x}\right\|_0 \leqslant s. \end{aligned} \tag{21}$$

Proof: It is obvious that x_s can be obtained by setting all but the largest (in magnitude) s components of x to zero. A solution to (21) can be obtained in the same way, which implies that x_s is also a solution to (21). ∎

Before proceeding further, the following lemma, which can be easily derived from the well-known Cauchy-Schwarz inequality, is presented as a useful tool.

Lemma 4: For any vector $z \in \mathbb{R}^m$, the following relation holds:

$$zz^T \leqslant z^T z I.$$

To analyze the characteristics of the decompression error, we propose the following theorem.

Theorem 1: At time instant k, we apply CoSaMP$_s$ after setting the number of iterations t_k according to (20). Then, with the decompression error defined in (9), d_k satisfies

$$\mathbb{E}\left\{d_k d_k^T\right\} \leqslant \xi_k \Psi^{-1} \left(\Psi^{-1}\right)^T,$$

where ξ_k is a scalar given as

$$\xi_k \triangleq 2\kappa^2 \left(1 + \frac{\sqrt{n}}{\sqrt{s}}\right)^2 \operatorname{tr}\left\{\Psi R_k \Psi^T\right\} + 2\left(\theta + \kappa\bar{\varepsilon}\right)^2 \tag{22}$$

Proof: Using (2) and (4), we have

$$z_k = \Psi h\left(x_k\right) + \Psi v_k. \tag{23}$$

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication.
Citation information: DOI 10.1109/TAC.2025.3589276, IEEE Transactions on Automatic Control

According to Assumption 1, the term $\Psi h(x_k)$ is a sparse vector with no more than s non-zero components, namely

$$\|\Psi h(x_k)\|_0 \leqslant s. \tag{24}$$

Define z_k^s as an s-sparse vector obtained by preserving s components with the largest magnitudes in z_k and setting all other components to zero. Following (15), z_k^s satisfies

$$\begin{aligned} z_k^s = &\ \underset{\bar{z}}{\operatorname{argmin}} \|z_k - \bar{z}\|_2 \\ &\text{s.t. } \|\bar{z}\|_0 \leqslant s. \end{aligned} \tag{25}$$

According to Lemma 3, the following relation holds as well:

$$\begin{aligned} z_k^s = &\ \underset{\bar{z}}{\operatorname{argmin}} \|z_k - \bar{z}\|_1 \\ &\text{s.t. } \|\bar{z}\|_0 \leqslant s. \end{aligned}$$

Due to the optimality of z_k^s, for any z satisfying $\|z\|_0 \leqslant s$, we have

$$\|z_k - z_k^s\|_2 \leqslant \|z_k - z\|_2, \tag{26}$$
$$\|z_k - z_k^s\|_1 \leqslant \|z_k - z\|_1. \tag{27}$$

Combining (23), (24), (26), and (27) yields

$$\|z_k - z_k^s\|_2 \leqslant \|z_k - \Psi h(x_k)\|_2 = \|\Psi v_k\|_2, \tag{28}$$
$$\|z_k - z_k^s\|_1 \leqslant \|z_k - \Psi h(x_k)\|_1 = \|\Psi v_k\|_1. \tag{29}$$

According to Lemma 2, (20), (28), and (29), we have

$$\begin{aligned} &\|z_k - \hat{z}_k\|_2 \\ &\leqslant \theta + \kappa\left(\|z_k - z_k^s\|_2 + \frac{1}{\sqrt{s}}\|z_k - z_k^s\|_1 + \|\varepsilon_k\|_2\right) \\ &\leqslant \theta + \kappa\left(\|\Psi v_k\|_2 + \frac{1}{\sqrt{s}}\|\Psi v_k\|_1 + \|\varepsilon_k\|_2\right) \\ &\leqslant \kappa\left(1 + \frac{\sqrt{n}}{\sqrt{s}}\right)\|\Psi v_k\|_2 + (\theta + \kappa\|\varepsilon_k\|_2). \end{aligned} \tag{30}$$

Considering the square on both sides of (30) leads to

$$\begin{aligned} &\|z_k - \hat{z}_k\|_2^2 \\ &= (z_k - \hat{z}_k)^T (z_k - \hat{z}_k) \\ &\leqslant \kappa^2\left(1 + \frac{\sqrt{n}}{\sqrt{s}}\right)^2 \|\Psi v_k\|_2^2 + (\theta + \kappa\|\varepsilon_k\|_2)^2 \\ &\quad + 2\kappa\left(1 + \frac{\sqrt{n}}{\sqrt{s}}\right)\|\Psi v_k\|_2 (\theta + \kappa\|\varepsilon_k\|_2) \\ &\leqslant 2\kappa^2\left(1 + \frac{\sqrt{n}}{\sqrt{s}}\right)^2 \|\Psi v_k\|_2^2 + 2(\theta + \kappa\|\varepsilon_k\|_2)^2 \\ &\leqslant 2\kappa^2\left(1 + \frac{\sqrt{n}}{\sqrt{s}}\right)^2 \operatorname{tr}\left\{\Psi v_k v_k^T \Psi^T\right\} + 2(\theta + \kappa\bar{\varepsilon})^2. \end{aligned} \tag{31}$$

Taking the mathematical expectation of both sides of (31) with respect to v_k, we have

$$\begin{aligned} &\mathbb{E}\left\{(z_k - \hat{z}_k)^T (z_k - \hat{z}_k)\right\} \\ &\leqslant 2\kappa^2\left(1 + \frac{\sqrt{n}}{\sqrt{s}}\right)^2 \mathbb{E}\left\{\operatorname{tr}\left\{\Psi v_k v_k^T \Psi^T\right\}\right\} + 2(\theta + \kappa\bar{\varepsilon})^2 \\ &\leqslant 2\kappa^2\left(1 + \frac{\sqrt{n}}{\sqrt{s}}\right)^2 \operatorname{tr}\left\{\Psi R_k \Psi^T\right\} + 2(\theta + \kappa\bar{\varepsilon})^2 = \xi_k \end{aligned} \tag{32}$$

It follows from Lemma 4 and (32) that

$$\begin{aligned} &\mathbb{E}\left\{(z_k - \hat{z}_k)(z_k - \hat{z}_k)^T\right\} \\ &\leqslant \mathbb{E}\left\{(z_k - \hat{z}_k)^T (z_k - \hat{z}_k) I\right\} \\ &= \mathbb{E}\left\{(z_k - \hat{z}_k)^T (z_k - \hat{z}_k)\right\} I = \xi_k I, \end{aligned}$$

where I represents the $m \times m$ identity matrix.

Following the definitions of z_k and $\breve{y}_k$, we have

$$\begin{aligned} &\mathbb{E}\left\{d_k d_k^T\right\} \\ &= \mathbb{E}\left\{\Psi^{-1}(z_k - \hat{z}_k)(z_k - \hat{z}_k)^T \left(\Psi^{-1}\right)^T\right\} \\ &= \Psi^{-1}\mathbb{E}\left\{(z_k - \hat{z}_k)(z_k - \hat{z}_k)^T\right\}\left(\Psi^{-1}\right)^T \\ &\leqslant \xi_k \Psi^{-1}\left(\Psi^{-1}\right)^T, \end{aligned}$$

and the proof is thus complete. ■

Remark 4: Theorem 1 points out that the decompression error of the measurement is bounded. According to (22), the upper bound approaches zero when the measurement noise covariance R_k, the upper bound $\bar{\varepsilon}$ of quantization error, and θ converge towards zero. The characteristic of decompression error holds significant importance as it reflects the impact of the compression-decompression process on state estimation.

B. Design of State Estimator

As stated in Theorem 1, an upper bound of $\mathbb{E}\left\{d_k d_k^T\right\}$ is known and can be further utilized in the state estimator design. Using (14), the update rule of the state estimator can be rewritten as

$$\hat{x}_{k+1|k+1} = \hat{x}_{k+1|k} + K_{k+1}\tilde{y}_{k+1|k} - K_{k+1}d_{k+1}.$$

In light of Theorem 1, it's acknowledged that only an upper bound of $\mathbb{E}\left\{d_{k+1}d_{k+1}^T\right\}$ is known, with the analytical expressions for the mean or covariance of d_{k+1} remaining elusive. This complexity poses substantial challenges in analytically determining the value of the estimation error covariance $P_{k+1|k+1}$. Consequently, it is very hard to obtain the optimal filter gain by minimizing $P_{k+1|k+1}$. As a workaround, a variance-constrained methodology is adopted to deduce an upper bound for $P_{k+1|k+1}$ and to formulate a suboptimal filter gain by minimizing this upper bound of $P_{k+1|k+1}$ [24].

To facilitate the derivation, a useful lemma is presented below.

Lemma 5 ([24]): For any real vectors x, y with the same dimension and positive scalar $\sigma > 0$, the following inequality holds:

$$xy^T + yx^T \leqslant \sigma xx^T + \sigma^{-1}yy^T.$$

The establishment of the upper bound for $P_{k+1|k+1}$ and the formulation of the suboptimal filter gain are presented in the subsequent theorem.

Theorem 2: For an arbitrary positive scalar $a > 0$, an upper bound of the estimation error covariance $P_{k+1|k+1}$ is calculated as:

$$\begin{aligned} \bar{P}_{k+1|k+1} \triangleq &\ (1+a)\Big(P_{k+1|k} - K_{k+1}P_{xy,k+1|k}^T \\ &- P_{xy,k+1|k}K_{k+1}^T + K_{k+1}P_{yy,k+1|k}K_{k+1}^T\Big) \end{aligned}$$

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication.
Citation information: DOI 10.1109/TAC.2025.3589276, IEEE Transactions on Automatic Control

$$+\left(1+2a^{-1}\right)\xi_{k+1}K_{k+1}\Psi^{-1}\left(\Psi^{-1}\right)^{T}K_{k+1}^{T}. \tag{33}$$

To minimize the trace of $\bar{P}_{k+1|k+1}$, the gain matrix K_{k+1} is computed as follows:

$$K_{k+1} = P_{xy,k+1|k}\left[P_{yy,k+1|k}+\frac{1+2a^{-1}}{1+a}\xi_{k+1}\Psi^{-1}\left(\Psi^{-1}\right)^{T}\right]^{-1}.$$

Proof: According to (12) and (13), we have

$$\tilde{x}_{k+1|k+1}=\tilde{x}_{k+1|k}-K_{k+1}\tilde{y}_{k+1|k}+K_{k+1}d_{k+1}. \tag{34}$$

Then, the estimation error covariance can be computed as:

$$\begin{aligned}
&P_{k+1|k+1}\\
&=\mathbb{E}\left\{\tilde{x}_{k+1|k+1}\tilde{x}_{k+1|k+1}^{T}\right\}\\
&=P_{k+1|k}-P_{xy,k+1|k}K_{k+1}^{T}-K_{k+1}P_{xy,k+1|k}^{T}\\
&\quad+\mathbb{E}\left\{\tilde{x}_{k+1|k}d_{k+1}^{T}\right\}K_{k+1}^{T}+K_{k+1}\mathbb{E}\left\{d_{k+1}\tilde{x}_{k+1|k}^{T}\right\}\\
&\quad+K_{k+1}P_{yy,k+1|k}K_{k+1}^{T}\\
&\quad-K_{k+1}\mathbb{E}\left\{\tilde{y}_{k+1|k}d_{k+1}^{T}\right\}K_{k+1}^{T}\\
&\quad-K_{k+1}\left\{d_{k+1}\tilde{y}_{k+1|k}^{T}\right\}K_{k+1}^{T}\\
&\quad+K_{k+1}\mathbb{E}\left\{d_{k+1}d_{k+1}^{T}\right\}K_{k+1}^{T}.
\end{aligned} \tag{35}$$

Before proceeding further, we give some more details about $\hat{y}_{k+1|k}$ according to the results in [26], [43]. Considering the linear expansion of $h(\cdot)$, $\tilde{y}_{k+1|k}$ can be written as:

$$\tilde{y}_{k+1|k}=\Gamma_{k+1}H_{k+1}\tilde{x}_{k+1|k}+v_{k+1}, \tag{36}$$

where

$$H_{k+1}\triangleq\left.\frac{\partial h(x)}{\partial x}\right|_{x=\hat{x}_{k+1|k}}$$

is a Jacobian matrix, and

$$\Gamma_{k+1}\triangleq\operatorname{diag}\left\{\gamma_{k+1,1},\gamma_{k+1,2},\cdots,\gamma_{k+1,m}\right\}$$

is applied to quantify the linearization errors. As will be shown later, Γ_{k+1} only serves as a placeholder here and we do not need to determine the specific value of Γ_{k+1} during state estimator design.

With the expansion of $\hat{y}_{k+1|k}$, we have

$$\begin{aligned}
P_{xy,k+1|k}&=\mathbb{E}\left\{\tilde{x}_{k+1|k}\left(\Gamma_{k+1}H_{k+1}\tilde{x}_{k+1|k}+v_{k+1}\right)^{T}\right\},\\
&=P_{k+1|k}H_{k+1}^{T}\Gamma_{k+1},
\end{aligned} \tag{37}$$

$$\begin{aligned}
P_{yy,k+1|k}&=\mathbb{E}\left\{\left(\Gamma_{k+1}H_{k+1}\tilde{x}_{k+1|k}+v_{k+1}\right)\right.\\
&\quad\left.\times\left(\Gamma_{k+1}H_{k+1}\tilde{x}_{k+1|k}+v_{k+1}\right)^{T}\right\}\\
&=\Gamma_{k+1}H_{k+1}P_{k+1|k}H_{k+1}^{T}\Gamma_{k+1}+R_{k+1}.
\end{aligned} \tag{38}$$

Substituting (37) and (38) into (35) yields

$$\begin{aligned}
&P_{k+1|k+1}\\
&=\left(I-K_{k+1}\Gamma_{k+1}H_{k+1}\right)P_{k+1|k}\left(I-K_{k+1}\Gamma_{k+1}H_{k+1}\right)^{T}\\
&\quad+\mathcal{S}_{k+1}+\mathcal{S}_{k+1}^{T}-\mathcal{T}_{k+1}-\mathcal{T}_{k+1}^{T}\\
&\quad+K_{k+1}R_{k+1}K_{k+1}^{T}+K_{k+1}\mathbb{E}\left\{d_{k+1}d_{k+1}^{T}\right\}K_{k+1}^{T},
\end{aligned} \tag{39}$$

where

$$\begin{aligned}
\mathcal{S}_{k+1}&\triangleq\mathbb{E}\left\{\left(I-K_{k+1}\Gamma_{k+1}H_{k+1}\right)\tilde{x}_{k+1|k}d_{k+1}^{T}K_{k+1}^{T}\right\},\\
\mathcal{T}_{k+1}&\triangleq\mathbb{E}\left\{K_{k+1}v_{k+1}d_{k+1}^{T}K_{k+1}^{T}\right\}.
\end{aligned}$$

According to Lemma 5, for any positive scalar $a>0$, we have

$$\begin{aligned}
\mathcal{S}_{k+1}+\mathcal{S}_{k+1}^{T}&\leqslant a\mathbb{E}\left\{\left(I-K_{k+1}\Gamma_{k+1}H_{k+1}\right)P_{k+1|k}\right.\\
&\quad\left.\times\left(I-K_{k+1}\Gamma_{k+1}H_{k+1}\right)^{T}\right\}\\
&\quad+a^{-1}K_{k+1}\mathbb{E}\left\{d_{k+1}d_{k+1}^{T}\right\}K_{k+1}^{T},
\end{aligned} \tag{40}$$

$$\begin{aligned}
-\mathcal{T}_{k+1}-\mathcal{T}_{k+1}^{T}&\leqslant aK_{k+1}R_{k+1}K_{k+1}^{T}+\\
&\quad+a^{-1}K_{k+1}\mathbb{E}\left\{d_{k+1}d_{k+1}^{T}\right\}K_{k+1}^{T}.
\end{aligned} \tag{41}$$

Using Theorem 1, we obtain the following inequality:

$$\mathbb{E}\left\{d_{k+1}d_{k+1}^{T}\right\}\leqslant\xi_{k+1}\Psi^{-1}\left(\Psi^{-1}\right)^{T}. \tag{42}$$

Combining (39), (40), (41), and (42), we have

$$\begin{aligned}
P_{k+1|k+1}\leqslant&\ (1+a)\left(I-K_{k+1}\Gamma_{k+1}H_{k+1}\right)P_{k+1|k}\\
&\times\left(I-K_{k+1}\Gamma_{k+1}H_{k+1}\right)^{T}\\
&+(1+a)K_{k+1}R_{k+1}K_{k+1}^{T}\\
&+\left(1+2a^{-1}\right)\xi_{k+1}K_{k+1}\Psi^{-1}\left(\Psi^{-1}\right)^{T}K_{k+1}^{T}\\
=&\ \bar{P}_{k+1|k+1}.
\end{aligned} \tag{43}$$

When designing the filter gain K_{k+1}, we aim to minimize the trace of the upper bound $\bar{P}_{k+1|k+1}$. Considering the partial derivatives of $\operatorname{tr}\left\{\bar{P}_{k+1|k+1}\right\}$, we have

$$\begin{aligned}
\frac{\partial\operatorname{tr}\left\{\bar{P}_{k+1|k+1}\right\}}{\partial K_{k+1}}=&\ 2(1+a)\left(-P_{k+1|k}H_{k+1}^{T}\Gamma_{k+1}\right.\\
&\left.+K_{k+1}\Gamma_{k+1}H_{k+1}P_{k+1|k}H_{k+1}^{T}\Gamma_{k+1}\right)\\
&+2(1+a)K_{k+1}R_{k+1}\\
&+2\left(1+2a^{-1}\right)\xi_{k+1}K_{k+1}\Psi^{-1}\left(\Psi^{-1}\right)^{T}.
\end{aligned}$$

By letting

$$\frac{\partial\operatorname{tr}\left\{\bar{P}_{k+1|k+1}\right\}}{\partial K_{k+1}}=0,$$

the filter gain K_{k+1} can be designed as

$$\begin{aligned}
K_{k+1}=&\ (1+a)P_{k+1|k}H_{k+1}^{T}\Gamma_{k+1}\left[(1+a)\Gamma_{k+1}\right.\\
&\times H_{k+1}P_{k+1|k}H_{k+1}^{T}\Gamma_{k+1}+(1+a)R_{k+1}\\
&\left.+\left(1+2a^{-1}\right)\xi_{k+1}\Psi^{-1}\left(\Psi^{-1}\right)^{T}\right]^{-1}\\
=&\ P_{k+1|k}H_{k+1}^{T}\Gamma_{k+1}\left(\Gamma_{k+1}H_{k+1}P_{k+1|k}H_{k+1}^{T}\Gamma_{k+1}\right.\\
&\left.+R_{k+1}+\frac{1+2a^{-1}}{1+a}\xi_{k+1}\Psi^{-1}\left(\Psi^{-1}\right)^{T}\right)^{-1},
\end{aligned} \tag{44}$$

Using (37) and (38), K_{k+1} can be rewritten as

$$K_{k+1}=P_{xy,k+1|k}\left(P_{yy,k+1|k}+\frac{1+2a^{-1}}{1+a}\xi_{k+1}\Psi^{-1}\left(\Psi^{-1}\right)^{T}\right)^{-1}. \tag{45}$$

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication.
Citation information: DOI 10.1109/TAC.2025.3589276, IEEE Transactions on Automatic Control

Noting that Γ_{k+1} is not involved in the expression of K_{k+1}, the specific value of Γ_{k+1} is not required. With (43) and (45) derived, the proof is complete. ■

Remark 5: In light of the derivation in Theorem 2, for any $a > 0$, $\bar{P}_{k+1|k+1}$ is an upper bound of the estimation error covariance and K_{k+1} can always minimize the trace of $\bar{P}_{k+1|k+1}$. The value of a will be determined later in (85) considering performance analysis.

Given that only upper bounds of estimation error covariances are attainable, it necessitates a modification to the UKF such that upper bounds are utilized in lieu of exact values when computing sigma points. In alignment with the approach detailed in [24], the execution of the modified UKF is encapsulated within Algorithm 2.

Remark 6: With the introduction of Algorithm 2, the issue of remote state estimation has been addressed for a specific category of nonlinear systems influenced by quantization effects and the process of data compression-decompression. Embedded within the UKF framework, the proposed state estimation algorithm offers a straightforward implementation.

C. Performance Analysis

In this part, we aim to conduct the performance analysis of the designed state estimator by assessing the boundedness of the estimation error. First, some useful lemmas are presented.

Lemma 6 ([20]): For arbitrary real vectors x, y, positive-definite matrix P with compatible dimensions, and positive scalar $\sigma > 0$, the following inequality holds:

$$x^T P y + y^T P x \leqslant \sigma x^T P x + \sigma^{-1} y^T P y.$$

Lemma 7 ([43]): For positive definite matrices $A > 0$, $C > 0$, and matrix B with compatible dimensions, the following inequality holds:

$$A^{-1} > B\left(B^T A B + C\right)^{-1} B^T.$$

Lemma 8: By defining

$$\tilde{R}_{k+1} \triangleq R_{k+1} + \frac{1+2a^{-1}}{1+a}\xi_{k+1}\Psi^{-1}\left(\Psi^{-1}\right)^T, \tag{46}$$

K_{k+1} and $\bar{P}_{k+1|k+1}$ given in Theorem 2 can be rewritten as follows:

$$K_{k+1} = \frac{1}{1+a}\bar{P}_{k+1|k+1}H_{k+1}^T\Gamma_{k+1}\tilde{R}_{k+1}^{-1}, \tag{47}$$

$$\bar{P}_{k+1|k+1}^{-1} = \frac{1}{1+a}\Big(P_{k+1|k}^{-1} + H_{k+1}^T\Gamma_{k+1}\tilde{R}_{k+1}^{-1}\Gamma_{k+1}H_{k+1}\Big). \tag{48}$$

Proof: Substituting (44) into (43), we have

$$\begin{aligned}\bar{P}_{k+1|k+1} =& (1+a)\left[P_{k+1|k} - P_{k+1|k}H_{k+1}^T\Gamma_{k+1}\right.\\ &\times\left(\Gamma_{k+1}H_{k+1}P_{k+1|k}H_{k+1}^T\Gamma_{k+1} + \tilde{R}_{k+1}\right)^{-1}\\ &\left.\times\Gamma_{k+1}H_{k+1}P_{k+1|k}\right]\\ =& (1+a)\left(I - K_{k+1}\Gamma_{k+1}H_{k+1}\right)P_{k+1|k}.\end{aligned} \tag{49}$$

For the sake of simplicity, define

$$M_{k+1} \triangleq \Gamma_{k+1}H_{k+1}P_{k+1|k}H_{k+1}^T\Gamma_{k+1} + \tilde{R}_{k+1}. \tag{50}$$

Algorithm 2 Modified UKF

Input: Initial state estimation $\hat{x}_{0|0}$, initial estimation error covariance $P_{0|0}$, and system model information.

1: Let $\bar{P}_{0|0} = P_{0|0}$ and $k = 0$.

2: At time instant k, generate $2n+1$ sigma points as follows:

$$\begin{aligned}\chi_{k|k}^0 &= \hat{x}_{k|k},\\ \chi_{k|k}^i &= \hat{x}_{k|k} + \left(\sqrt{(\lambda+n)\,\bar{P}_{k|k}}\right)_i, 1 \leqslant i \leqslant n,\\ \chi_{k|k}^i &= \hat{x}_{k|k} - \left(\sqrt{(\lambda+n)\,\bar{P}_{k|k}}\right)_{i-n}, n+1 \leqslant i \leqslant 2n,\end{aligned}$$

where λ is a scaling factor defined in [39].

3: Obtain predictions of sigma points as

$$\chi_{k+1|k}^i = f\left(\chi_{k|k}^i, u_k\right), 0 \leqslant i \leqslant 2n.$$

4: Compute the prediction and the prediction covariance as

$$\begin{aligned}\hat{x}_{k+1|k} &= \sum_{i=0}^{2n}\mathcal{W}_i^m\chi_{k+1|k}^i,\\ P_{k+1|k} &= \sum_{i=0}^{2n}\mathcal{W}_i^c\left(\chi_{k+1|k}^i - \hat{x}_{k+1|k}\right)\\ &\quad\times\left(\chi_{k+1|k}^i - \hat{x}_{k+1|k}\right)^T + Q_k,\end{aligned}$$

where $\mathcal{W}_i^m$ and $\mathcal{W}_i^c$ are constant weights defined in [39].

5: Compute the measurement prediction and covariance matrices as follows:

$$\begin{aligned}\mathcal{Y}_{k+1|k}^i &= g\left(\chi_{k+1|k}^i\right), 0 \leqslant i \leqslant 2n,\\ \hat{y}_{k+1|k} &= \sum_{i=0}^{2n}\mathcal{W}_i^m\mathcal{Y}_{k+1|k}^i,\\ P_{yy,k+1|k} &= \sum_{i=0}^{2n}\mathcal{W}_i^c\left(\mathcal{Y}_{k+1|k}^i - \hat{y}_{k+1|k}\right)\\ &\quad\times\left(\mathcal{Y}_{k+1|k}^i - \hat{y}_{k+1|k}\right)^T + R_{k+1},\\ P_{xy,k+1|k} &= \sum_{i=0}^{2n}\mathcal{W}_i^c\left(\chi_{k+1|k}^i - \hat{x}_{k+1|k}\right)\\ &\quad\times\left(\mathcal{Y}_{k+1|k}^i - \hat{y}_{k+1|k}\right)^T.\end{aligned}$$

6: Determine the filter gain K_{k+1} according to (45).

7: Update the state estimation $\hat{x}_{k+1|k+1}$ according to (11).

8: Compute $\bar{P}_{k+1|k+1}$ according to (33).

9: Let $k \leftarrow k+1$ and return to Step 2 until the maximum step is reached.

Combining (44), (46), and (50) yields

$$\begin{aligned}K_{k+1} &= P_{k+1|k}H_{k+1}^T\Gamma_{k+1}M_{k+1}^{-1}\\ &= P_{k+1|k}H_{k+1}^T\Gamma_{k+1}\tilde{R}_{k+1}^{-1}\\ &\quad - P_{k+1|k}H_{k+1}^T\Gamma_{k+1}M_{k+1}^{-1}M_{k+1}\tilde{R}_{k+1}^{-1}\\ &\quad + P_{k+1|k}H_{k+1}^T\Gamma_{k+1}M_{k+1}^{-1}\tilde{R}_{k+1}\tilde{R}_{k+1}^{-1}\\ &= \left(I - P_{k+1|k}H_{k+1}^T\Gamma_{k+1}M_{k+1}^{-1}\Gamma_{k+1}H_{k+1}\right)\end{aligned}$$

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication.
Citation information: DOI 10.1109/TAC.2025.3589276, IEEE Transactions on Automatic Control

$$\begin{aligned}&\times P_{k+1|k}H_{k+1}^T\Gamma_{k+1}\tilde{R}_{k+1}^{-1}\\ =&\left(I-K_{k+1}\Gamma_{k+1}H_{k+1}\right)P_{k+1|k}H_{k+1}^T\Gamma_{k+1}\tilde{R}_{k+1}^{-1}.\end{aligned}\tag{51}$$

According to (49) and (51), (47) can be obtained. Besides, applying the matrix inversion lemma to (49), we have

$$\begin{aligned}\bar{P}_{k+1|k+1}=&\left(1+a\right)\Big(P_{k+1|k}^{-1}\\&+H_{k+1}^T\Gamma_{k+1}\tilde{R}_{k+1}^{-1}\Gamma_{k+1}H_{k+1}\Big)^{-1},\end{aligned}\tag{52}$$

which indicates that (48) holds. Thus, the proof is complete. ■

Before proceeding to the performance analysis, we consider the linear expansion of $\tilde{x}_{k+1|k}$ as follows:

$$\tilde{x}_{k+1|k}=\Lambda_k F_k\tilde{x}_{k|k}+w_k,\tag{53}$$

where the Jacobian matrix F_k is defined as

$$F_k\triangleq\left.\frac{\partial f\left(x,u\right)}{\partial x}\right|_{x=\hat{x}_{k|k},u=u_k},$$

and

$$\Lambda_k\triangleq\operatorname{diag}\left\{\lambda_{k,1},\lambda_{k,2},\cdots,\lambda_{k,n}\right\}$$

quantifies the linearization errors. Then, we make the following assumption about system parameters:

Assumption 2: There exist positive constants $f_{\min}$, $f_{\max}$, $h_{\min}$, $h_{\max}$, $\lambda_{\min}$, $\lambda_{\max}$, $\gamma_{\min}$, $\gamma_{\max}$, $q_{\min}$, $q_{\max}$, $r_{\max}$, and $\psi_{\max}$ such that the following bounds hold for any $k>0$:

$$f_{\min}^2I\leqslant F_kF_k^T\leqslant f_{\max}^2I,\tag{54}$$
$$h_{\min}^2I\leqslant H_kH_k^T\leqslant h_{\max}^2I,\tag{55}$$
$$\lambda_{\min}^2I\leqslant\Lambda_k\Lambda_k^T\leqslant\lambda_{\max}^2I,\tag{56}$$
$$\gamma_{\min}^2I\leqslant\Gamma_k\Gamma_k^T\leqslant\gamma_{\max}^2I.\tag{57}$$
$$q_{\min}I\leqslant Q_k\leqslant q_{\max}I,\tag{58}$$
$$r_{\min}I\leqslant R_k\leqslant r_{\max}I,\tag{59}$$
$$\psi_{\min}I\leqslant\Psi\leqslant\psi_{\max}I,\tag{60}$$
$$\xi_{\min}\leqslant\xi_k\leqslant\xi_{\max},\tag{61}$$

Define

$$\tilde{r}_{\min}\triangleq r_{\min}+\frac{1+2a^{-1}}{1+a}\frac{\xi_{\min}}{\psi_{\max}^2},$$
$$\tilde{r}_{\max}\triangleq r_{\max}+\frac{1+2a^{-1}}{1+a}\frac{\xi_{\max}}{\psi_{\min}^2},$$

we have the following theorem.

Theorem 3: Under Assumption 2, let $\bar{\eta}$ be a positive scalar such that

$$\bar{P}_{0|0}\leqslant\bar{\eta}I,\tag{62}$$
$$\frac{1}{\left(1+a\right)\bar{\eta}}\leqslant\frac{1}{\lambda_{\max}^2f_{\max}^2\bar{\eta}+q_{\max}}+\frac{h_{\min}^2\gamma_{\min}^2}{\tilde{r}_{\max}}.\tag{63}$$

Then, there exist positive constants $p_{\min}$ and $p_{\max}$ such that

$$p_{\min}I\leqslant\bar{P}_{k|k}\leqslant p_{\max}I.$$

Proof: Using (53), we have

$$P_{k+1|k}=\Lambda_kF_k\bar{P}F_k^T\Lambda_k+Q_k.\tag{64}$$

It follows from (48) and (64) that

$$\begin{aligned}\bar{P}_{k+1|k+1}^{-1}=&\frac{1}{1+a}\Big[\left(\Lambda_kF_k\bar{P}F_k^T\Lambda_k+Q_k\right)^{-1}\\&+H_{k+1}^T\Gamma_{k+1}\tilde{R}_{k+1}^{-1}\Gamma_{k+1}H_{k+1}\Big]\\\leqslant&\frac{1}{1+a}\left(Q_k^{-1}+H_{k+1}^T\Gamma_{k+1}\tilde{R}_{k+1}^{-1}\Gamma_{k+1}H_{k+1}\right).\end{aligned}$$

According to (59)–(61) and (46), $\tilde{R}_{k+1}$ is bounded by

$$\tilde{r}_{\min}I\leqslant\tilde{R}_{k+1}\leqslant\tilde{r}_{\max}.\tag{65}$$

Using Assumption 2 and (65), we have

$$\bar{P}_{k+1|k+1}^{-1}\leqslant\frac{1}{1+a}\left(\frac{1}{q_{\min}}+\frac{h_{\max}^2\gamma_{\max}^2}{\tilde{r}_{\min}}\right)I.$$

Constructing

$$p_{\min}\triangleq\left[\frac{1}{1+a}\left(\frac{1}{q_{\min}}+\frac{h_{\max}^2\gamma_{\max}^2}{\tilde{r}_{\min}}\right)\right]^{-1},$$

we have

$$\bar{P}_{k+1|k+1}\geqslant p_{\min}I.\tag{66}$$

Next, let us prove the existence of the upper bound by mathematical induction. Since $\bar{P}_{0|0}\leqslant\bar{\eta}I$ is bounded, we assume $\bar{P}_{k|k}\leqslant\bar{\eta}I$. Using (48), (63), and (64), we have

$$\begin{aligned}\bar{P}_{k+1|k+1}\leqslant&\left(1+a\right)\left[\frac{1}{\lambda_{\max}^2f_{\max}^2\bar{\eta}+q_{\max}}+\frac{h_{\min}^2\gamma_{\min}^2}{\tilde{r}_{\max}}\right]^{-1}I\\\leqslant&\bar{\eta}I.\end{aligned}$$

Then, construct

$$p_{\max}\triangleq\bar{\eta},$$

and we have

$$\bar{P}_{k+1|k+1}\leqslant p_{\max}I.\tag{67}$$

It follows from (66) and (67) that the proof is complete. ■

Remark 7: The boundedness of $\bar{P}_{k+1|k+1}$ is intricately linked to the observability and detectability of the system [18], [35]. Given the focus on nonlinear systems within this paper, pinpointing precise conditions for observability or detectability proves to be impractical. As a feasible alternative, a sufficient condition for the boundedness of the estimation error will be contemplated. Theorem 3 serves as an important foundation of Theorem 4, which is a core result of performance analysis proposing conditions for the mean-square exponential boundedness of the estimation error.

With the boundedness of $\bar{P}_{k+1|k+1}$ now assured, attention shifts towards examining the properties of the estimation error $\tilde{x}_{k+1|k+1}$. To facilitate this analysis, the subsequent lemma is presented as a foundational instrument.

Lemma 9 ([1], [37]): For a stochastic process η_k, assume that there exist a stochastic process $V\left(\eta_k\right)$ and constants $\nu_{\min}>0$, $\nu_{\max}>0$, $\mu>0$, and $0<\lambda\leqslant1$ such that the following two conditions are fulfilled for any $k>0$:

$$\nu_{\min}\left\|\eta_k\right\|_2^2\leqslant V\left(\eta_k\right)\leqslant\nu_{\max}\left\|\eta_k\right\|_2^2,$$
$$\mathbb{E}\left\{V\left(\eta_k\right)\mid\eta_{k-1}\right\}-V\left(\eta_{k-1}\right)\leqslant-\lambda V\left(\eta_{k-1}\right)+\mu.$$

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication.
Citation information: DOI 10.1109/TAC.2025.3589276, IEEE Transactions on Automatic Control

Then, the stochastic process η_k is exponentially bounded in mean square, namely

$$\mathbb{E}\left\{\|\eta_k\|_2^2\right\} \leqslant \frac{\nu_{\max}}{\nu_{\min}}\mathbb{E}\left\{\|\eta_0\|_2^2\right\}(1-\lambda)^k + \frac{\mu}{\nu_{\min}}\sum_{i=1}^{k-1}(1-\lambda)^i.$$

With the lemmas and intermediate results mentioned above, the main results for performance analysis are presented as follows.

Theorem 4: Consider the modified UKF designed in this paper. Under Assumption 2, if the conditions in Theorem 3 are satisfied, then there exists a constant $a > 0$ such that the estimation error $\tilde{x}_{k|k}$ is exponentially bounded in mean square.

Proof: We start the proof by constructing

$$V_k\left(\tilde{x}_{k|k}\right) \triangleq \tilde{x}_{k|k}^T \bar{P}_{k|k}^{-1} \tilde{x}_{k|k}.$$

According to Theorem 3, $V_k\left(\tilde{x}_{k|k}\right)$ is bounded by

$$\frac{1}{p_{\max}}\left\|\tilde{x}_{k|k}\right\|_2^2 \leqslant V_k\left(\tilde{x}_{k|k}\right) \leqslant \frac{1}{p_{\min}}\left\|\tilde{x}_{k|k}\right\|_2^2. \tag{68}$$

Using (34), (36), and (48), we have

$$\begin{aligned}
&\mathbb{E}\left\{V_{k+1}\left(\tilde{x}_{k+1|k+1}\right) \mid \tilde{x}_{k|k}\right\} \\
&= \mathbb{E}\left\{\left[\left(I - K_{k+1}\Gamma_{k+1}H_{k+1}\right)\tilde{x}_{k+1|k} - K_{k+1}v_{k+1} \right.\right. \\
&\quad \left. + K_{k+1}d_{k+1}\right]^T \bar{P}_{k+1|k+1}^{-1}\left[\left(I - K_{k+1}\Gamma_{k+1}H_{k+1}\right)\tilde{x}_{k+1|k}\right. \\
&\quad \left.\left. - K_{k+1}v_{k+1} + K_{k+1}d_{k+1}\right] \mid \tilde{x}_{k|k}\right\}.
\end{aligned}$$

According to Lemma 6, for any $\sigma > 0$, we can obtain

$$\begin{aligned}
&\mathbb{E}\left\{V_{k+1}\left(\tilde{x}_{k+1|k+1}\right) \mid \tilde{x}_{k|k}\right\} \\
&\leqslant (1+\sigma)\,\mathbb{E}\left\{\left[\left(I - K_{k+1}\Gamma_{k+1}H_{k+1}\right)\tilde{x}_{k+1|k}\right.\right. \\
&\quad - K_{k+1}v_{k+1}\Big]^T \bar{P}_{k+1|k+1}^{-1}\left[\left(I - K_{k+1}\Gamma_{k+1}H_{k+1}\right)\tilde{x}_{k+1|k}\right. \\
&\quad \left.\left. - K_{k+1}v_{k+1}\right] \mid \tilde{x}_{k|k}\right\} + \left(1+\sigma^{-1}\right) \\
&\quad \times \mathbb{E}\left\{\left(K_{k+1}d_{k+1}\right)^T \bar{P}_{k+1|k+1}^{-1}\left(K_{k+1}d_{k+1}\right) \mid \tilde{x}_{k|k}\right\}.
\end{aligned} \tag{69}$$

Defining

$$\begin{aligned}
\Xi_{k+1} \triangleq\ & \left[\left(I - K_{k+1}\Gamma_{k+1}H_{k+1}\right)\tilde{x}_{k+1|k} - K_{k+1}v_{k+1}\right]^T \\
& \times \bar{P}_{k+1|k+1}^{-1}\left[\left(I - K_{k+1}\Gamma_{k+1}H_{k+1}\right)\tilde{x}_{k+1|k}\right. \\
& \left. - K_{k+1}v_{k+1}\right],
\end{aligned} \tag{70}$$

we have

$$\begin{aligned}
&\Xi_{k+1} \\
&= \tilde{x}_{k+1|k}^T \bar{P}_{k+1|k+1}^{-1}\tilde{x}_{k+1|k} + \left(\Gamma_{k+1}H_{k+1}\tilde{x}_{k+1|k} + v_{k+1}\right)^T \\
&\quad \times K_{k+1}^T \bar{P}_{k+1|k+1}^{-1} K_{k+1}\left(\Gamma_{k+1}H_{k+1}\tilde{x}_{k+1|k} + v_{k+1}\right) \\
&\quad - \tilde{x}_{k+1|k}^T \bar{P}_{k+1|k+1}^{-1} K_{k+1}\left(\Gamma_{k+1}H_{k+1}\tilde{x}_{k+1|k} + v_{k+1}\right) \\
&\quad - \left(\Gamma_{k+1}H_{k+1}\tilde{x}_{k+1|k} + v_{k+1}\right)^T K_{k+1}^T \bar{P}_{k+1|k+1}^{-1}\tilde{x}_{k+1|k}.
\end{aligned} \tag{71}$$

Substituting (44) and (48) into (71) yields

$$\begin{aligned}
\Xi_{k+1} =\ & \frac{1}{1+a}\tilde{x}_{k+1|k}^T P_{k+1|k}^{-1}\tilde{x}_{k+1|k} \\
& - \left(\Gamma_{k+1}H_{k+1}\tilde{x}_{k+1|k}\right)^T\left(\frac{1}{1+a}\tilde{R}_{k+1}^{-1}\right. \\
& - \frac{1}{(1+a)^2}\tilde{R}_{k+1}^{-1}\Gamma_{k+1}H_{k+1}\bar{P}_{k+1|k+1}H_{k+1}^T \\
& \left.\times\Gamma_{k+1}\tilde{R}_{k+1}^{-1}\right)\left(\Gamma_{k+1}H_{k+1}\tilde{x}_{k+1|k}\right) \\
& + \frac{1}{(1+a)^2}v_{k+1}^T\tilde{R}_{k+1}^{-1}\Gamma_{k+1}H_{k+1}\bar{P}_{k+1|k+1} \\
& \times H_{k+1}^T\Gamma_{k+1}\tilde{R}_{k+1}^{-1}v_{k+1}.
\end{aligned} \tag{72}$$

Let

$$\begin{aligned}
\Pi_{k+1} \triangleq\ & \frac{1}{1+a}\tilde{R}_{k+1}^{-1} - \frac{1}{(1+a)^2}\tilde{R}_{k+1}^{-1}\Gamma_{k+1}H_{k+1} \\
& \times \bar{P}_{k+1|k+1}H_{k+1}^T\Gamma_{k+1}\tilde{R}_{k+1}^{-1}.
\end{aligned}$$

According to (44), (46), and (47), we have

$$\begin{aligned}
&\Pi_{k+1} \\
&= \frac{1}{1+a}\tilde{R}_{k+1}^{-1}\left(I - \Gamma_{k+1}H_{k+1}K_{k+1}\right) \\
&= \frac{1}{1+a}\tilde{R}_{k+1}^{-1}\left[I - \Gamma_{k+1}H_{k+1}P_{k+1|k}H_{k+1}^T\Gamma_{k+1}\right. \\
&\quad \left.\times\left(\Gamma_{k+1}H_{k+1}P_{k+1|k}H_{k+1}^T\Gamma_{k+1} + \tilde{R}_{k+1}^{-1}\right)^{-1}\right] \\
&= \frac{1}{1+a}\left(\Gamma_{k+1}H_{k+1}P_{k+1|k}H_{k+1}^T\Gamma_{k+1} + \tilde{R}_{k+1}\right)^{-1}.
\end{aligned} \tag{73}$$

Combining (72) and (73) leads to

$$\begin{aligned}
&\Xi_{k+1} \\
&= \frac{1}{1+a}\left[\tilde{x}_{k+1|k}^T P_{k+1|k}^{-1}\tilde{x}_{k+1|k} - \left(\Gamma_{k+1}H_{k+1}\tilde{x}_{k+1|k}\right)^T\right. \\
&\quad \times\left(\Gamma_{k+1}H_{k+1}P_{k+1|k}H_{k+1}^T\Gamma_{k+1} + \tilde{R}_{k+1}\right)^{-1} \\
&\quad \left.\times\left(\Gamma_{k+1}H_{k+1}\tilde{x}_{k+1|k}\right)\right] + \frac{1}{(1+a)^2}v_{k+1}^T\tilde{R}_{k+1}^{-1}\Gamma_{k+1} \\
&\quad \times H_{k+1}\bar{P}_{k+1|k+1}H_{k+1}^T\Gamma_{k+1}\tilde{R}_{k+1}^{-1}v_{k+1}.
\end{aligned}$$

Since w_k is Gaussian white noise, we have

$$\begin{aligned}
&\mathbb{E}\left\{\Xi_{k+1} \mid \tilde{x}_{k|k}\right\} \\
&= \mathbb{E}\left\{\frac{1}{1+a}\left[\left(\Lambda_k F_k\tilde{x}_{k|k}\right)^T P_{k+1|k}^{-1}\left(\Lambda_k F_k\tilde{x}_{k|k}\right)\right.\right. \\
&\quad + w_k^T P_{k+1|k}^{-1}w_k - \left(\Gamma_{k+1}H_{k+1}\Lambda_k F_k\tilde{x}_{k|k}\right)^T N_{k+1}^{-1} \\
&\quad \times\left(\Gamma_{k+1}H_{k+1}\Lambda_k F_k\tilde{x}_{k|k}\right) - \left(\Gamma_{k+1}H_{k+1}w_k\right)^T \\
&\quad \times N_{k+1}^{-1}\left(\Gamma_{k+1}H_{k+1}w_k\right)\Big] + \frac{1}{(1+a)^2}v_{k+1}^T\tilde{R}_{k+1}^{-1} \\
&\quad \left.\left.\times\Gamma_{k+1}H_{k+1}\bar{P}_{k+1|k+1}H_{k+1}^T\Gamma_{k+1}\tilde{R}_{k+1}^{-1}v_{k+1} \mid \tilde{x}_{k|k}\right\}.\right.
\end{aligned} \tag{74}$$

where

$$\begin{aligned}
N_{k+1} \triangleq\ & \Gamma_{k+1}H_{k+1}\Lambda_k F_k\bar{P}_{k|k}F_k^T\Lambda_k H_{k+1}^T\Gamma_{k+1} \\
& + \Gamma_{k+1}H_{k+1}Q_k H_{k+1}\Gamma_{k+1} + \tilde{R}_{k+1}.
\end{aligned}$$

Inequalities (54) and (56) imply that $\left(\Lambda_k F_k\right)^{-1}$ exists. Therefore, by applying Lemma 7 and (64), we acquire that

$$\begin{aligned}
&\left(\Lambda_k F_k\tilde{x}_{k|k}\right)^T P_{k+1|k}^{-1}\left(\Lambda_k F_k\tilde{x}_{k|k}\right) \\
&\leqslant \tilde{x}_{k|k}^T\left(\Lambda_k F_k\right)^T\left(\Lambda_k F_k\bar{P}_{k|k}F_k^T\Lambda_k\right)^{-1}\left(\Lambda_k F_k\right)\tilde{x}_{k|k}
\end{aligned}$$

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication.
Citation information: DOI 10.1109/TAC.2025.3589276, IEEE Transactions on Automatic Control

$$= \tilde{x}_{k|k}^T \bar{P}_{k|k}^{-1} \tilde{x}_{k|k} = V_k\left(\tilde{x}_{k|k}\right). \quad (75)$$

According to Lemma 7, we have

$$\left(\Gamma_{k+1}H_{k+1}\Lambda_k F_k\right)^T N_{k+1}^{-1}\left(\Gamma_{k+1}H_{k+1}\Lambda_k F_k\right) < \bar{P}_{k|k}^{-1}. \quad (76)$$

Defining

$$\zeta_{k+1} \triangleq \frac{\tilde{x}_{k|k}^T \left(\Gamma_{k+1}H_{k+1}\Lambda_k F_k\right)^T N_{k+1}^{-1}\left(\Gamma_{k+1}H_{k+1}\Lambda_k F_k\right)\tilde{x}_{k|k}}{\tilde{x}_{k|k}^T \bar{P}_{k|k}^{-1}\tilde{x}_{k|k}},$$

it follows from (76) that $\zeta_{k+1} < 1$.

Using (54)–(57), (65), and Theorem 3, we have

$$\begin{aligned}\zeta_{k+1} &\geqslant p_{\min}\left(\gamma_{\min}h_{\min}\lambda_{\min}f_{\min}\right)^2\left[q_{\max}\gamma_{\max}^2 h_{\max}^2\right.\\ &\quad\left.+\tilde{r}_{\max}+p_{\max}\left(\gamma_{\max}h_{\max}\lambda_{\max}f_{\max}\right)^2\right]\\ &\triangleq \zeta_{\min}.\end{aligned}$$

From the definitions of ζ_{k+1} and $\zeta_{\min}$, we obtain

$$\begin{aligned}&\tilde{x}_{k|k}^T\left(\Gamma_{k+1}H_{k+1}\Lambda_k F_k\right)^T N_{k+1}^{-1}\left(\Gamma_{k+1}H_{k+1}\Lambda_k F_k\right)\tilde{x}_{k|k}\\ &\geqslant \zeta_{\min}V_k\left(\tilde{x}_{k|k}\right). \quad (77)\end{aligned}$$

Additionally, define

$$\begin{aligned}\mu_{k+1} &\triangleq \frac{1}{1+a}\Big[w_k^T P_{k+1|k}^{-1}w_k - \left(\Gamma_{k+1}H_{k+1}w_k\right)^T\\ &\quad\times N_{k+1}^{-1}\left(\Gamma_{k+1}H_{k+1}w_k\right)\Big] + \frac{1}{(1+a)^2}v_{k+1}^T\\ &\quad\times \tilde{R}_{k+1}^{-1}\Gamma_{k+1}H_{k+1}\bar{P}_{k+1|k+1}H_{k+1}^T\Gamma_{k+1}\tilde{R}_{k+1}^{-1}v_{k+1}. \quad (78)\end{aligned}$$

Applying Lemma 7 yields

$$\begin{aligned}&\left(\Gamma_{k+1}H_{k+1}w_k\right)^T N_{k+1}^{-1}\left(\Gamma_{k+1}H_{k+1}w_k\right)\\ &< w_k^T\left(\Lambda_k F_k\bar{P}_{k|k}F_k^T\Lambda_k + Q_k\right)^{-1}w_k\\ &= w_k^T P_{k+1|k}^{-1}w_k.\end{aligned}$$

Subsequently, we have

$$\mu_{k+1} > 0.$$

Since μ_{k+1} is a scalar, using the property of trace, the upper bound of μ_{k+1} can be obtained as

$$\begin{aligned}\mu_{k+1} &= \operatorname{tr}\left\{\mu_{k+1}\right\}\\ &\leqslant \frac{1}{1+a}\operatorname{tr}\left\{\left(\Lambda_k F_k\bar{P}_{k|k}F_k^T\Lambda_k + Q_k\right)^{-1}Q_k\right\}\\ &\quad+\frac{1}{(1+a)^2}\operatorname{tr}\left\{\tilde{R}_{k+1}^{-1}\Gamma_{k+1}H_{k+1}\bar{P}_{k+1|k+1}\right.\\ &\quad\left.\times H_{k+1}^T\Gamma_{k+1}\tilde{R}_{k+1}^{-1}R_{k+1}\right\}\\ &\leqslant \frac{1}{1+a}n + \frac{1}{(1+a)^2}\frac{\gamma_{\max}^2 h_{\max}^2 p_{\max}r_{\max}}{\tilde{r}_{\min}^2}m\\ &\triangleq \mu_{\max}. \quad (79)\end{aligned}$$

Combining (74), (75), (77), and (78), we have

$$\mathbb{E}\left\{\Xi_{k+1} \mid \tilde{x}_{k|k}\right\} \leqslant \frac{1-\zeta_{\min}}{1+a}V_k\left(\tilde{x}_{k|k}\right) + \mathbb{E}\left\{\mu_{k+1} \mid \tilde{x}_{k|k}\right\}. \quad (80)$$

Let

$$\begin{aligned}\varrho_{k+1} &\triangleq (1+\sigma)\,\mathbb{E}\left\{\mu_{k+1} \mid \tilde{x}_{k|k}\right\} + \left(1+\sigma^{-1}\right)\\ &\quad\times\mathbb{E}\left\{\left(K_{k+1}d_{k+1}\right)^T\bar{P}_{k+1|k+1}^{-1}\left(K_{k+1}d_{k+1}\right) \mid \tilde{x}_{k|k}\right\}\\ &> 0, \quad (81)\end{aligned}$$

and

$$\begin{aligned}\bar{p}_{\min} &\triangleq q_{\min} + \lambda_{\min}^2 f_{\min}^2 p_{\min},\\ \bar{p}_{\max} &\triangleq q_{\max} + \lambda_{\max}^2 f_{\max}^2 p_{\max},\\ k_{\max} &\triangleq \frac{\bar{p}_{\max}h_{\max}\gamma_{\max}}{\bar{p}_{\min}h_{\min}^2\gamma_{\min}^2 + \tilde{r}_{\min}^2}.\end{aligned}$$

Then, it follows from (44) and (64) that

$$\bar{p}_{\min}I \leqslant P_{k+1|k} \leqslant \bar{p}_{\max}, \quad K_{k+1}K_{k+1}^T \leqslant k_{\max}^2 I.$$

Subsequently, we have

$$\begin{aligned}&\left(K_{k+1}d_{k+1}\right)^T\bar{P}_{k+1|k+1}^{-1}\left(K_{k+1}d_{k+1}\right)\\ &= \operatorname{tr}\left\{\xi_{k+1}K_{k+1}^T\bar{P}_{k+1|k+1}^{-1}K_{k+1}\Psi^{-1}\left(\Psi^{-1}\right)^T\right\}\\ &\leqslant \frac{k_{\max}^2\xi_{\max}}{p_{\min}\psi_{\min}^2}m \triangleq \varsigma_{\max}. \quad (82)\end{aligned}$$

According to (79), (81), and (82), an upper bound of ϱ_{k+1} can be expressed as

$$\varrho_{k+1} \leqslant (1+\sigma)\,\mu_{\max} + \left(1+\sigma^{-1}\right)\varsigma_{\max} \triangleq \varrho_{\max}. \quad (83)$$

Define

$$\phi \triangleq 1 - \frac{(1+\sigma)(1-\zeta_{\min})}{1+a}.$$

From (69), (70), (80), and (83), it is illustrated that

$$\begin{aligned}&\mathbb{E}\left\{V_{k+1}\left(\tilde{x}_{k+1|k+1}\right) \mid \tilde{x}_{k|k}\right\} - V_k\left(\tilde{x}_{k|k}\right)\\ &\leqslant -\phi V_k\left(\tilde{x}_{k|k}\right) + \varrho_{\max}. \quad (84)\end{aligned}$$

Since $0 < \zeta_{\min} < 1$, we can choose a and σ satisfying:

$$0 < a < \frac{1-\zeta_{\min}}{\zeta_{\min}}, \quad 0 < \sigma < \frac{\zeta_{\min}}{1-\zeta_{\min}}. \quad (85)$$

Then, it follows that

$$0 < \phi < 1. \quad (86)$$

Since (68), (84), and (86) hold, we can apply Lemma 9. Therefore, $\tilde{x}_{k|k}$ is exponentially bounded in mean square and satisfies

$$\begin{aligned}\mathbb{E}\left\{\left\|\tilde{x}_{k|k}\right\|_2^2\right\} &\leqslant \frac{p_{\max}}{p_{\min}}\mathbb{E}\left\{\left\|x_0\right\|_2^2\right\}(1-\phi)^k\\ &\quad+\frac{\varrho_{\max}}{p_{\min}}\sum_{i=1}^{k-1}(1-\phi)^i,\end{aligned}$$

which completes the proof. ■

Remark 8: Theorem 4 suggests that, despite the presence of measurement noise and quantization errors, the state estimator employing decompressed data maintains assured performance. The convergence rate of the developed estimation algorithm is related to ϕ. Specifically, the convergence is quick when ϕ is close to 1. This theorem addresses the previously identified

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication.
Citation information: DOI 10.1109/TAC.2025.3589276, IEEE Transactions on Automatic Control

gap, where scant research has delved into the impact of CS-based data compression and decompression on the performance of state estimation.

Remark 9: This paper distinguishes itself from the existing literature on state estimation in NCSs by introducing novel contributions that center around the integration of CS and analysis for handling high-dimensional and redundant measurement data, particularly in the presence of measurement noise and quantization errors. The key distinctions and novel contributions of this work can be summarized as follows:

1) *Innovative Use of CS*: While CS has been explored in various contexts, this paper uniquely applies CS to the compression and decompression of measurement data in NCSs.
2) *Comprehensive Analysis of Decompression Error*: The paper provides a detailed examination of the decompression error introduced by CS, factoring in both measurement noise and quantization errors.
3) *Development of a Modified UKF*: The adaptation of the UKF to accommodate the nuances of CS-based data compression and decompression, particularly through the innovative handling of decompression error, showcases a novel integration of CS with established state estimation techniques.
4) *Rigorous Performance Analysis*: The paper goes beyond simple implementation by offering a thorough performance analysis that includes the derivation of a sufficient condition for the boundedness of estimation error.
5) *Practical Implications and Versatility*: By demonstrating the effectiveness of the proposed approach through simulation experiments on power grid systems (in the next section) and potentially other large-scale systems, the paper underscores the practical applicability and versatility of its contributions.

IV. Numerical Simulations

In this paper, numerical simulations are conducted on the IEEE 69 bus system to assess the performance of the proposed state estimation method. The system architecture of the IEEE 69 bus system is shown in Fig. 3. As stated in [28], power grid systems are typical examples of systems whose measurement vectors can be sparsified. In this study, we follow the structure of the power system given in [32], and the system configuration described in [5].

When performing simulations, we follow the system model utilized in [28]. The state variable is defined as:

$$x_k \triangleq \begin{bmatrix} |V_1|_k & \cdots & |V_N|_k & \theta_{1,k} & \cdots & \theta_{N,k} \end{bmatrix}^T,$$

where $N = 69$ is the total number of buses. $|V_i|_k$ and $\theta_{i,k}$ $(i \in \{1, 2, \cdots, N\})$ represent the voltage magnitude and the voltage angle of the i-th bus at time instant k, respectively. By tradition, the first bus is set as the slack bus serving as a reference for all other buses, and we have $|V_1|_k = 1$ and $\theta_{1,k} = 0$ for any $k > 0$. In this paper, we mainly focus on the tracking problem of the power grid and assume the system state is slowly varying over time. Therefore, we consider the following state transition equation:

$$x_{k+1} = x_k + w_k,$$

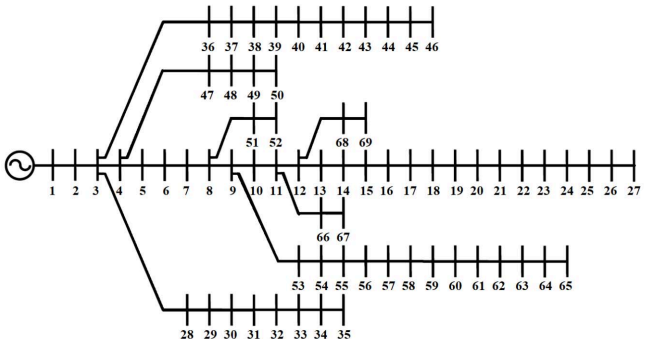


Fig. 3: System architecture of the IEEE 69 bus system [38]

where w_k is the white Gaussian process noise with known covariance matrix $Q_k = 10^{-4}I$ [28]. Besides, $P_{0|0}$ is set as $10^{-2}I$ and a is chosen as 0.1.

The measurement of the considered power grid system consists of the active and reactive injected power of each bus. As to the i-th bus at time instant k, its active injected power $P_{i,k}$ and reactive injected power $Q_{i,k}$ can be computed using the following nonlinear expressions [28]:

$$P_{i,k} = |V_i|_k \sum_{j=1}^{N} |V_j|_k \left(G_{ij}\cos\theta_{ij,k} + B_{ij}\sin\theta_{ij,k}\right), \quad (87)$$

$$Q_{i,k} = |V_i|_k \sum_{j=1}^{N} |V_j|_k \left(G_{ij}\sin\theta_{ij,k} - B_{ij}\cos\theta_{ij,k}\right), \quad (88)$$

where $\theta_{ij,k} \triangleq \theta_{i,k} - \theta_{j,k}$ denotes the difference between the voltage angles of the i-th and the j-th bus. Known constants G_{ij} and B_{ij} represent the real and imaginary terms of the element at the i-th row and j-th column of the power grid admittance matrix, which is dependent on the configuration of the power grid system. Finally, the measurement vector is defined as

$$y_k \triangleq \begin{bmatrix} P_{1,k} & \cdots & P_{N,k} & Q_{1,k} & \cdots & Q_{N,k} \end{bmatrix}^T + v_k, \quad (89)$$

where v_k indicates the zero-mean Gaussian measurements with known covariance $R_k = 10^{-4}I$.

To design Φ, we follow [8] and set Φ as a random matrix of which each entry is drawn from i.i.d. Gaussian distribution $\mathcal{N}(0, 1/p)$ with $p = 40$. By evaluating the measurements of the IEEE 69 bus system, we found that the Haar matrix [27], [28] performs well in obtaining sparse representations of measurement vectors. Therefore, we apply the Haar matrix as the sparsifying transformation. Then, the sparsified measurement can be compressed and decompressed by the CS method. Fig. 4 illustrates the original measurement and the decompressed measurement. Here, the measurement is split into two parts, namely active injected power and reactive injected power. It is shown that the values of active injected power are relatively large and strongly correlated, and can thus be decompressed accurately. In contrast, the values of the reactive injected power are small and prone to measurement noise and quantization errors. As a result, the difference between the original reactive injected power and the decompressed values is relatively large.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication.
Citation information: DOI 10.1109/TAC.2025.3589276, IEEE Transactions on Automatic Control

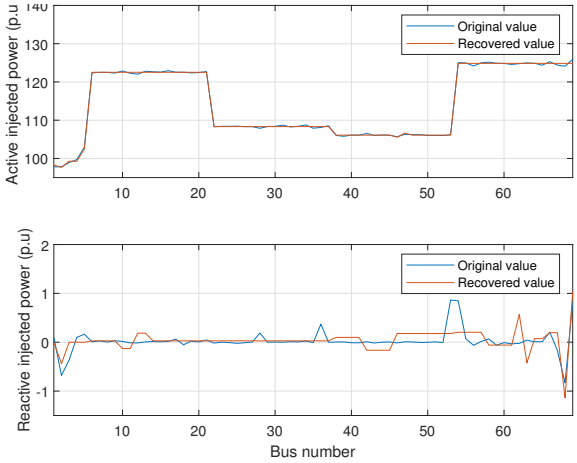


Fig. 4: Original and decompressed measurement

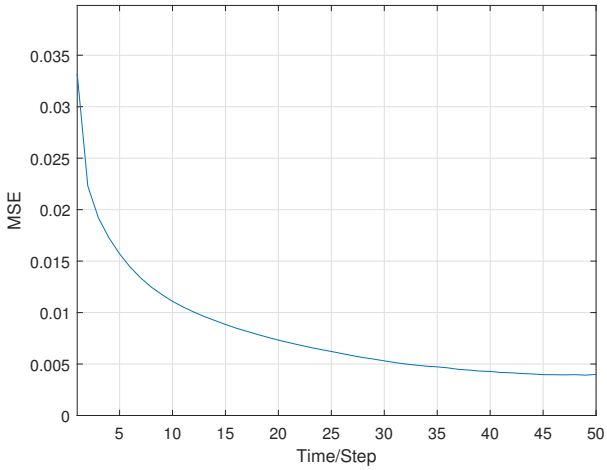


Fig. 6: MSE of state estimation

Fig. 5 illustrates the true value and estimated value of the state vector at the last time instant. It can be seen that both the voltage magnitude and voltage angle can be precisely estimated. Fig. 6 shows the mean-square error (MSE) of state estimation over time. It is demonstrated that the designed state estimator can quickly converge and reach a high estimation performance.

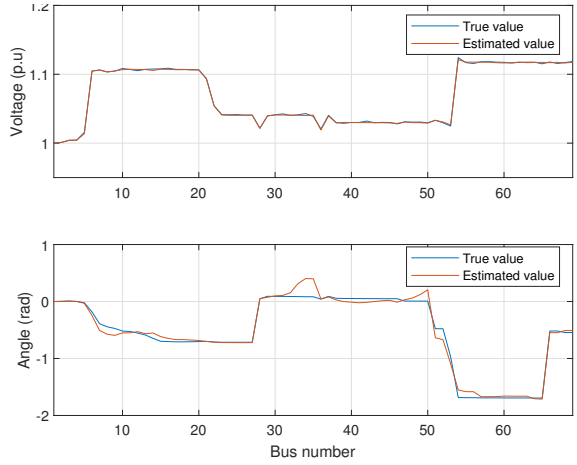


Fig. 5: State estimation at the last time instant

To verify the necessity of data decompression, we follow the intuitive method described in Remark 2, which can be operated without data decompression. As shown in Fig. 7, the method without data decompression can hardly obtain accurate state estimation. Furthermore, without data decompression, the MSE is much higher than the MSE of our method with data decompression. Therefore, data decompression is a crucial step that effectively utilizes information of sparsify and maintains the observability of the system.

Since it is infeasible to analytically obtain the expression of $\mathbb{E}\left\{d_k d_k^T\right\}$, an upper bound of $\mathbb{E}\left\{d_k d_k^T\right\}$ is derived in Theorem 1 as an alternative. The utilization of Lemma 1 and inequality manipulations inevitably lead to conservativeness in the upper bound of $\mathbb{E}\left\{d_k d_k^T\right\}$, and thus increases the conservativeness in the upper bound of estimation error covariance.

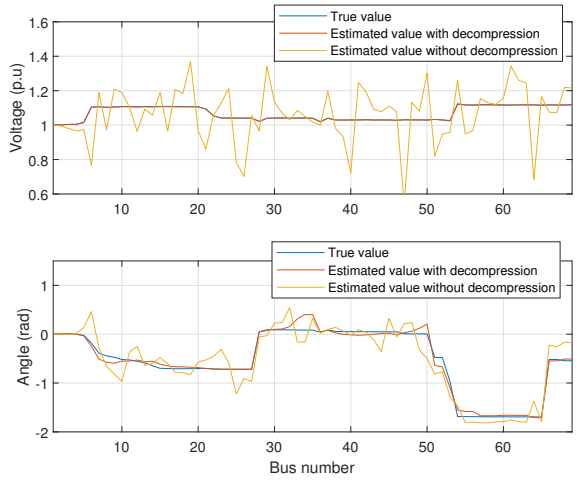


Fig. 7: Comparison of state estimation at the last time instant

To reduce the conservativeness, one possible countermeasure is to perform Monte Carlo experiments to assess the statistics of the vector $z_k - \hat{z}_k$ and subsequently provide an estimate of $\mathbb{E}\left\{d_k d_k^T\right\}$. In other words, this method focuses on obtaining an empirical upper bound of $\mathbb{E}\left\{d_k d_k^T\right\}$ to assist the practical implementation of the designed approach.

V. Conclusions

In this paper, the remote state estimation has been addressed for a class of NCSs integrated with data compression-decompression. Considering the redundancy embedded in the measurements, CS has been utilized for high-performance data compression-decompression. Then, a modified UKF has been applied for the state estimation task using decompressed measurement data. A key challenge of this task is that measurement noise and quantization errors lead to nonnegligible decompression errors, which inevitably degrade the state estimation accuracy. Accordingly, the combined impact of measurement noise and quantization errors on decompression errors has been analyzed. Then, filter gains of the modified

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication.
Citation information: DOI 10.1109/TAC.2025.3589276, IEEE Transactions on Automatic Control

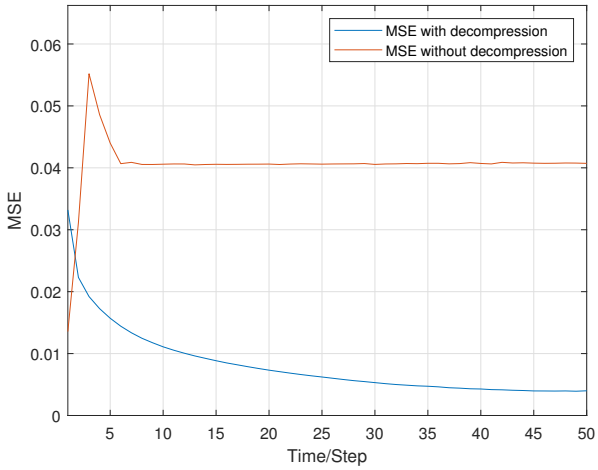


Fig. 8: Comparison of MSE

UKF have been devised to minimize an upper bound of the estimation error covariance. A sufficient condition has also been given to ensure that the estimation error is exponentially bounded in the mean square. Finally, simulation experiments have been conducted based on the IEEE 69 bus system to verify the effectiveness of the proposed state estimation method.

REFERENCES

[1] R. G. Agniel and E. I. Jury, Almost sure boundedness of randomly sampled systems, *SIAM Journal on Control*, vol. 9, no. 3, pp. 372–384, 1971.

[2] A. Ajorloo, A. Amini, E. Tohidi, M. H. Bastani, and G. Leus, Antenna placement in a compressive sensing-based colocated MIMO radar, *IEEE Transactions on Aerospace and Electronic Systems*, vol. 56, no. 6, pp. 4606–4614, 2020.

[3] S. M. S. Alam, B. Natarajan, and A. Pahwa, Distribution grid state estimation from compressed measurements, *IEEE Transactions on Smart Grid*, vol. 5, no. 4, pp. 1631–1642, 2014.

[4] C. N. Babu, P. Sure, and C. M. Bhuma, Sparse Bayesian learning assisted approaches for road network traffic state estimation, *IEEE Transactions on Intelligent Transportation Systems*, vol. 22, no. 3, pp. 1733–1741, 2021.

[5] M. E. Baran and F. F. Wu, Network reconfiguration in distribution systems for loss reduction and load balancing, *IEEE Transactions on Power Delivery*, vol. 4, no. 2, pp. 1401–1407, 1989.

[6] T. Blumensath and M. E. Davies, Iterative hard thresholding for compressed sensing, *Applied and Computational Harmonic Analysis*, vol. 27, pp. 265–274, 2009.

[7] R. Caballero-Águila and J. Linares-Pérez, Distributed fusion filtering for uncertain systems with coupled noises, random delays and packet loss prediction compensation, *International Journal of Systems Science*, vol. 54, no. 2, pp. 371–390, 2023.

[8] E. J. Candes and T. Tao, Near-optimal signal recovery from random projections: universal encoding strategies?, *IEEE Transactions on Information Theory*, vol. 52, no. 12, pp. 5406–5425, 2006.

[9] M. Carlsson, D. Gerosa, and C. Olsson, An unbiased approach to compressed sensing, *Inverse Problems*, vol. 36, no. 11, art. no. 115014, 2020.

[10] D. Carta, C. Muscas, P. A. Pegoraro, A. V. Solinas, and S. Sulis, Compressive sensing-based harmonic sources identification in smart grids, *IEEE Transactions on Instrumentation and Measurement*, vol. 70, art. no. 9000810, 2020.

[11] X. Cheng, M. Wang, and Y. Guan, Ultrawideband channel estimation: a Bayesian compressive sensing strategy based on statistical sparsity, *IEEE Transactions on Vehicular Technology*, vol. 64, no. 5, pp. 1819–1832, 2015.

[12] D. Dai, J. Li, Y. Song, and F. Yang, Event-based recursive filtering for nonlinear bias-corrupted systems with amplify-and-forward relays, *Systems Science & Control Engineering*, vol. 12, no. 1, art. no. 2332419, 2024.

[13] D. L. Donoho, Compressed sensing, *IEEE Transactions on Information Theory*, vol. 52, no. 4, pp. 1289–1306, 2006.

[14] Y. Dua, V. Kumar, and R. S. Singh, Comprehensive review of hyperspectral image compression algorithms, *Optical Engineering*, vol. 59, no. 9, art. no. 090902, 2020.

[15] M. G. El-Mashed, Compressed sensing Kalman filter estimation scheme for MIMO system under phase noise problem, *IET Communications*, vol. 14, no. 22, pp. 4108–4115, 2020.

[16] Y. Gao, J. Hu, K. Chi, C. Jia, and J. Qi, A variance-constrained method to encoding-decoding H_∞ state estimation for memristive neural networks with energy harvesting sensor, *Neurocomputing*, vol. 579, art. no. 127448, 2024.

[17] D. Gurve, D. Delisle-Rodriguez, T. Bastos-Filho, and S. Krishnan, Trends in compressive sensing for EEG signal processing applications, *Sensors*, vol. 20, no. 13, art. no. 3703, 2020.

[18] M. Haring and T. A. Johansen, On the stability bounds of Kalman filters for linear deterministic discrete-time systems, *IEEE Transactions on Automatic Control*, vol. 65, no. 10, pp. 4434–4439, 2020.

[19] J. Hespanha, A. Ortega, and L. Vasudevan, Towards the control of linear systems with minimum bit-rate, *Proc. 15th Int. Symp. on Mathematical Theory of Networks and Systems (MTNS)*, 2002.

[20] N. Hou, Z. Wang, H. Dong, J. Hu, and X. Liu, Sensor fault estimation for nonlinear complex networks with time delays under saturated innovations: A binary encoding scheme, *IEEE Transactions on Network Science and Engineering*, vol. 9, no. 6, pp. 4171–4183, 2022.

[21] U. Jayasankar, V. Thirumal, and D. Ponnurangam, A survey on data compression techniques: from the perspective of data quality, coding schemes, data type and applications, *Journal of King Saud University - Computer and Information Sciences*, vol. 33, no. 2, pp. 119–140, 2021.

[22] X. Kan, Y. Fan, Z. Fang, L. Cao, N. N. Xiong, D. Yang, and X. Li, A novel IoT network intrusion detection approach based on Adaptive Particle Swarm Optimization Convolutional Neural Network, *Information Sciences*, vol. 568, pp. 147–162, 2021.

[23] H. S. Karimi and B. Natarajan, Kalman filtered compressive sensing with intermittent observations, *Signal Processing*, vol. 163, pp. 49–58, 2019.

[24] C. Li, Z. Wang, W. Song, S. Zhao, J. Wang, and J. Shan, Resilient unscented Kalman filtering fusion with dynamic event-triggered scheme: Applications to multiple unmanned aerial vehicles, *IEEE Transactions on Control Systems Technology*, vol. 31, no. 1, pp. 370–381, 2023.

[25] L. Li, Y. Fang, L. Liu, H. Peng, J. Kurths, and Y. Yang, Overview of compressed sensing: sensing model, reconstruction algorithm, and its applications, *Applied Sciences*, vol. 10, no. 17, art. no. 5909, 2020.

[26] Y. Liu, Z. Wang, and D. Zhou, UKF-based remote state estimation for discrete artificial neural networks with communication bandwidth constraints, *Neural Networks*, vol. 108, pp. 393–398, 2018.

[27] S. Mallat, *A wavelet tour of signal processing*. Elsevier, 1999.

[28] R. Mohammadrezaee, J. Ghaisari, G. Yousefi, and M. Kamali, Dynamic state estimation of smart distribution grids using compressed measurements, *IEEE Transactions on Smart Grid*, vol. 12, no. 5, pp. 4535–4542, 2021.

[29] D. Needell and J. A. Tropp, CoSaMP: iterative signal recovery from incomplete and inaccurate samples, *Applied and Computational Harmonic Analysis*, vol. 26, no. 3, pp. 301–321, 2009.

[30] D. Pan, String stable bidirectional platooning control for heterogeneous connected automated vehicles, *International Journal of Network Dynamics and Intelligence*, vol. 3, no. 4, art. no. 100026, 2024.

[31] Z.-H. Pang, L.-Z. Fan, J. Sun, K. Liu and G.-P. Liu, Detection of stealthy false data injection attacks against networked control systems via active data modification, *Information Sciences*, vol. 546, pp. 192–205, 2021.

[32] I. Pisica, C. Bulac, and M. Eremia, Optimal distributed generation location and sizing using genetic algorithms, *2009 15th International Conference on Intelligent System Applications to Power Systems*, 2009.

[33] B. Qu, D. Peng, Y. Shen, L. Zou, and B. Shen, A survey on recent advances on dynamic state estimation for power systems, *International Journal of Systems Science*, vol. 55, no. 16, pp. 3305–3321, 2024.

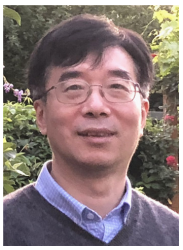
[34] F. Qu, E. Tian, and X. Zhao, Chance-constrained H_∞ state estimation for recursive neural networks under deception attacks and energy constraints: The finite-horizon case, *IEEE Transactions on Neural Networks and Learning Systems*, vol. 34, no. 9, pp. 6492–6503, 2023.

[35] K. Reif, S. Gunther, E. Yaz, and R. Unbehauen, Stochastic stability of the discrete-time extended Kalman filter, *IEEE Transactions on Automatic Control*, vol. 44, no. 4, pp. 714–728, 1999.

[36] W. Song, Z. Wang, Z. Li, and Q.-L. Han, Particle-filter-based state estimation for delayed artificial neural networks: when probabilistic saturation constraints meet redundant channels, *IEEE Transactions on Neural Networks and Learning Systems*, vol. 35, no. 3, pp. 4354–4362, 2024.
[37] T. J. Tarn and Y. Rasis, Observers for nonlinear stochastic systems, *IEEE Transactions on Automatic Control*, vol. 21, no. 4, pp. 441–448, 1976.
[38] C. Venkatesan, R. Kannadasan, M. H. Alsharif, M.-K. Kim, and J. Nebhen, A novel multiobjective hybrid technique for siting and sizing of distributed generation and capacitor banks in radial distribution systems, *Sustainability*, vol. 13, no. 6, art. no. 3308, 2021.
[39] E. A. Wan and R. Van Der Merwe, The unscented Kalman filter for nonlinear estimation, *Proceedings of the IEEE 2000 Adaptive Systems for Signal Processing, Communications, and Control Symposium*, 2000.
[40] X. Wan, Y. Li, Y. Li, and M. Wu, Finite-time H_∞ state estimation for two-time-scale complex networks under stochastic communication protocol, *IEEE Transactions on Neural Networks and Learning Systems*, vol. 33, no. 1, pp. 25-36, 2022.
[41] S. Wang, Z. Wang, H. Dong, Y. Chen, and G. Lu, Dynamic event-triggered quadratic nonfragile filtering for non-Gaussian systems: tackling multiplicative noises and missing measurements, *IEEE/CAA Journal of Automatica Sinica*, vol. 11, no. 5, pp. 1127–1138, 2024.
[42] Y.-A. Wang, B. Shen, L. Zou and Q.-L. Han, A survey on recent advances in distributed filtering over sensor networks subject to communication constraints, *International Journal of Network Dynamics and Intelligence*, vol. 2, no. 2, art. no. 100007, 2023.
[43] K. Xiong, H. Y. Zhang, and C. W. Chan, Performance evaluation of UKF-based nonlinear filtering, *Automatica*, vol. 42, no. 2, pp. 261–270, 2006.
[44] T. Xue and Y. Liu, State estimation based on measurement from a part of nodes for a delayed Itô type of complex network with Markovian mode-dependent parameters, *Neurocomputing*, vol. 600, art. no. 128188, 2024.
[45] C. Yang, J. Liang, and X. Chen, Distributed event-based H_∞ consensus filtering for 2-D T-S fuzzy systems over sensor networks subject to DoS attacks, *Information Sciences*, vol. 641, art. no. 119079, 2023.
[46] Y. Yuan, X. Tang, W. Zhou, W. Pan, X. Li, H.-T. Zhang, H. Ding and J. Goncalves, Data driven discovery of cyber physical systems, *Nature Communications*, vol. 10, no. 1, pp. 1–9, 2019.
[47] R. R. Zebari, A. M. Abdulazeez, D. Q. Zeebaree, D. A. Zebari, and J. N. Saeed, A comprehensive review of dimensionality reduction techniques for feature selection and feature extraction, *Journal of Applied Science and Technology Trends*, vol. 1, no. 2, pp. 56–70, 2020.
[48] J. Zhang, K. You, and L. Xie, Innovation compression for communication-efficient distributed optimization with linear convergence, *IEEE Transactions on Automatic Control*, vol. 68, no. 11, pp. 6899–6906, 2023.
[49] R. Zhang, H. Liu, Y. Liu and H. Tan, Dynamic event-triggered state estimation for discrete-time delayed switched neural networks with constrained bit rate, *Systems Science & Control Engineering*, vol. 12, no. 1, art. no. 2334304, 2024.

Jiahao Song received the B. Eng. degree in automation in 2020 from Tsinghua University, Beijing, China.

He is currently a Ph.D. candidate in control science and engineering at the Department of Automation, Tsinghua University, Beijing, China. From Feb. 2023 to Jan. 2025, he was an international researcher in the Department of Computer Science, Brunel University London, Uxbridge, United Kingdom. His research interests include state estimation, networked systems, compressive sensing, fault diagnosis, and machine learning.



Zidong Wang (Fellow, IEEE) received the B.Sc. degree in mathematics in 1986 from Suzhou University, Suzhou, China, and the M.Sc. degree in applied mathematics in 1990 and the Ph.D. degree in electrical engineering in 1994, both from Nanjing University of Science and Technology, Nanjing, China.

He is currently Professor of Dynamical Systems and Computing in the Department of Computer Science, Brunel University London, U.K. From 1990 to 2002, he held teaching and research appointments in universities in China, Germany and the U.K. Prof. Wang's research interests include dynamical systems, signal processing, bioinformatics, control theory and applications. He has published a number of papers in international journals. He is a holder of the Alexander von Humboldt Research Fellowship of Germany, the JSPS Research Fellowship of Japan, William Mong Visiting Research Fellowship of Hong Kong.

Prof. Wang serves (or has served) as the Editor-in-Chief for *International Journal of Systems Science*, the Editor-in-Chief for *Neurocomputing*, the Editor-in-Chief for *Systems Science & Control Engineering*, and an Associate Editor for 12 international journals including IEEE TRANSACTIONS ON AUTOMATIC CONTROL, IEEE TRANSACTIONS ON CONTROL SYSTEMS TECHNOLOGY, IEEE TRANSACTIONS ON NEURAL NETWORKS, IEEE TRANSACTIONS ON SIGNAL PROCESSING, and IEEE TRANSACTIONS ON SYSTEMS, MAN, AND CYBERNETICS-PART C. He is a Member of the Academia Europaea, a Member of the European Academy of Sciences and Arts, an Academician of the International Academy for Systems and Cybernetic Sciences, a Fellow of the IEEE, a Fellow of the Royal Statistical Society and a member of program committee for many international conferences.

Qinyuan Liu received the B.Eng. degree in measurement and control technology and instrumentation from Huazhong University of Science and Technology, Wuhan, China, in 2012, and the Ph.D. degree in control science and engineering from Tsinghua University, Beijing, China, in 2017.

He is currently deputy dean and a professor in the Department of Computer Science and Technology, Tongji University, Shanghai, China. From Jul. 2015 to Sep. 2016, he was a Research Assistant in the Department of Electronic & Computer Engineering, Hong Kong University of Science and Technology, Hong Kong, China. From Jan. 2016 to Jan. 2017, he was an international researcher in the Department of Computer Science, Brunel University London, UK. His research interests include networked control systems, multi-agent systems, and distributed filtering. He is an active reviewer for many international journals.

Xiao He (M'12-SM'20) received the B.Eng. degree in information technology from the Beijing Institute of Technology, Beijing, China, in 2004, and the Ph.D. degree in control science and engineering from Tsinghua University, Beijing, China, in 2010.

He is currently a Tenured Professor with the Department of Automation, Tsinghua University. He is also the Head of the Research Center of Safety Control, Tsinghua University. He has authored more than 200 papers in refereed international journals. His research interests include fault diagnosis and fault-tolerant control of dynamic systems, networked systems, cyber-physical systems, as well as their applications.

Dr. He is a Full Member of Sigma Xi, the Scientific Research Society, and a Senior Member of the Chinese Association of Automation. He is an Associate Editor for IEEE Transactions on Neural Networks and Learning Systems, IEEE Transactions on Automation Science and Engineering, and Control Engineering Practice.