ASTES

Advances in Science, Technology and Engineering Systems Journal
Vol. 7, No. 5, 146-156 (2022)
www.astesj.com
Special Issue on Innovation in Computing, Engineering Science & Technology

ASTES Journal
ISSN: 2415-6698

Model Order Reduction and Distribution for Efficient State Estimation in Sensor and Actuator Networks

Ferdinand Friedrich*, Christoph Ament

University of Augsburg, Chair of Control Theory, Augsburg, 86159, Germany

ARTICLE INFO

Article history:
Received: 31 August, 2022
Accepted: 04 October, 2022
Online: 25 October, 2022

Keywords:
Model Order Reduction
Sensor and Actuator Network
Distributed State Estimation

ABSTRACT

We present in this contribution the distribution of a global multi-input-multi-output system in a sensor and actuator network. Based on controllability and observability, the global system is decentralized and the system properties are preserved as a result. This results in multiple decentralized local single-input-single-output systems with the same system order as the global system. As these local systems are implemented on decentralized CPUs in the network, the computational effort of the nodes has to be minimized. This is achieved by approximating the input and output behavior and reducing the system order of the decentralized local systems. For this purpose, the two most common techniques, Balanced Truncation and Krylov subspace methods, are presented. Kalman filters are used for state reconstruction. To approximate the input/output behavior of the global system, information from all decentralized reduced local systems is necessary, thus a fully interconnected network is used for communication. By decentralized fusion algorithms in the network nodes, the Kalman filter algorithm is separated and distributed in the network.

1 Introduction

This paper is an extension of "Reduced and Distributed Estimation in Sensor and Actuator Networks - Automated Design Based on Controllability and Observability" presented at the IEEE Conference on Control Technology and Applications (CCTA) 2021 in San Diego (virtual) and "Model-Order Reduction and System Distribution Using Krylov Subspaces - An Approach for Efficient State Estimation in Sensor and Actuator Networks" presented at the IEEE Conference on Control Technology and Applications (CCTA) 2022 in Trieste [1], [2].

The requirements for the accuracy of mathematical models are constantly increasing. Also large-scale MIMO[1] systems are not protected from the increasing requirements. These and the increasing complexity of the systems enhance the system order of the dynamic large-scale MIMO systems [3], [4]. The use of such complex and large-scale models in simulation is very cumbersome and sometimes impossible. Moreover, system analysis and controller design using known methods is also impossible. For this reason, methods for linear and nonlinear systems are developed to reduce large models to a manageable size and to approximate the characteristic properties related to the input/output behavior [5]. Memory and computational requirements are reduced by distributing the system so that only parts of the model are available for local systems [6]. However, this requires decentralized estimation procedures.

If the internal model relationships (e.g. mechanical and electrical) are known, the distribution into local systems is performed manually [6]. Furthermore, system digraphs and cut-point sets are used for model distribution [7]. In [8], a manual method for distributing large systems is also presented. In this method, reduced local systems with overlapping state vectors are created. All described manual methods share the property that the actual system order of the global system is preserved in the network despite the distribution into local systems.

In contrast, we use decentralization for distribution based on the controllability and observability of the global MIMO system. By decoupling the inputs and outputs, several local SISO[2] systems are formed from the global MIMO system. These decentralized local systems are implemented in a fully interconnected network. Moreover, we use model order reduction techniques, resulting in the decentralized local systems with much lower order and preserving the system properties [9]. Further, the computational effort in each network node is reduced.

Several techniques for reducing the model order are known [10]. The most common techniques for approximating the input and output behavior are Balanced Truncation and *Krylov* subspace methods.

*Corresponding Author: Ferdinand Friedrich, University of Augsburg, Chair of Control Theory, Universitaetsstrasse 2, 86159 Augsburg, ferdinand.friedrich@uni-a.de
[1]Multi-Input-Multi-Output
[2]Single-Input-Single-Output

In Balanced Truncation, the Hankel singular values are analyzed and the states with a small impact on the energy transport are neglected and truncated [11]. In *Krylov* subspace methods, to approximate the input and output behavior, the transfer functions of the global and reduced systems are expanded at an arbitrary development point. In this process, the first coefficients, also called moments, of the Taylor series of the transfer function are matched [12], [13].

Therefore, the methods are also called moment matching with the *Padé*-type approximation method based on rational interpolation [14], [15]. If only a single development point (several are also possible) $s = 0$ is chosen, the reduction is ascribed to an *padé*-approximation [16]. A more extensive background on *Krylov* subspace methods is given in [10], [17], and [18]. Model order reduction with *Krylov* subspaces requires less memory and computational effort (compared to Balanced Truncation) [19]. Meanwhile, several algorithms for the Krylov subspaces are known. In this paper, we only use the *Arnoldi* algorithm with its extension to MIMO systems [20], [21].

This paper is organized as follows. In Section 2, the systems under consideration are defined. The methods used to reduce the model order are presented in Section 3. The distribution and deployment in a network are demonstrated in Section 4. A demonstrator is used for detailed real-time experimental evaluation of the presented methods and the results are presented in Section 5.

2 Preliminaries

In this contribution, we refer to global multi-input-multi-output linear time-invariant systems of the form

$$\mathbf{\Sigma}_G = \begin{cases} \dot{\mathbf{x}}(t) = \mathbf{A}\mathbf{x}(t) + \mathbf{B}\mathbf{u}(t) \\ \mathbf{y}(t) = \mathbf{C}\mathbf{x}(t) \end{cases} \tag{1}$$

Here, $\mathbf{A} \in \mathbb{R}^{n\times n}$ is the system matrix and $\mathbf{B} \in \mathbb{R}^{n\times p}$ is the input matrix. The matrix $\mathbf{C} \in \mathbb{R}^{q\times n}$ ist the output matrix and $\mathbf{x}$, $\mathbf{u}$ as well as $\mathbf{y}$ are the state vector, the inputs and the outputs of the global continuous-time system $\mathbf{\Sigma}_G$.

3 Model Order Reduction

Precise modeling of technical systems often leads to large systems of differential equations, requiring a high computational effort. For this reason, methods are used to reduce the order of the systems. Thereby the characteristic properties are supposed to be preserved and the input/output behavior has to be maintained as exactly as possible. The linear reduction methods can be categorized into three classes. One class is the modal reduction methods, in which less dominant eigenvalues and eigenmodes are removed from the system representation. The second class of reduction methods, which is also used in this contribution, are the balanced methods. Here, energy-oriented analyses are used to remove an unimportant subsystem. The third class describes the Krylov subspace methods. In this method, certain parameters of the transfer function, the so-called moments, are preserved in the reduced model. This class is also used for decentralization and distribution in this contribution.

3.1 Balanced Truncation

The Balanced Truncation (BT) method is based on the system properties of controllability and observability. Different methods for determining controllability and observability are known. For balanced truncation, the *Gramian* controllability matrix and the *Gramian* observability matrix shall be used.
A system (1) is controllable if a control vector $\mathbf{u}(t)$ exists that transforms the system from an initial state $\mathbf{x}(t = 0) = \mathbf{0}$ to a final state in finite time. The controllability *Gramian* is defined as

$$\mathbf{W}_c = \int_0^\infty e^{\mathbf{A}t}\,\mathbf{B}\mathbf{B}^T\,e^{\mathbf{A}^T t}\,dt. \tag{2}$$

Furthermore, a system (1) is fully observable if the system state $\mathbf{x}(t = 0)$ can be determined by observing the output $\mathbf{y}(t)$ over a finite time interval. The *Gramian* observability matrix is derived as follows

$$\mathbf{W}_o = \int_0^\infty e^{\mathbf{A}^T t}\,\mathbf{C}^T\mathbf{C}\,e^{\mathbf{A}t}\,dt. \tag{3}$$

For the transfer of the system (1) from the initial position $\mathbf{x}_0$ to the final position $\mathbf{x}_e$ in infinite time, the energy $\mathbf{x}_e^T\mathbf{W}_c^{-1}\mathbf{x}_e$ must be applied. In contrast, the energy $\mathbf{x}_0^T\mathbf{W}_o\mathbf{x}_0$ is generated when the system (1) in the initial position $\mathbf{x}_0$ can oscillate uninfluenced ($\mathbf{u}(t) = 0$) to its rest position. This energy-based approach reveals that a difficultly controllable state requires a high level of energy to achieve.This can also be observed in the eigenvalues of the matrix $\mathbf{W}_c^{-1}$. The more energy is needed to reach a specific state, the larger is its eigenvalue. This analysis can also be used for observability. However, here a small eigenvalue of the state determines that it is badly observable, because only a small amount of energy is visible.
By a state transformation the system $\mathbf{\Sigma}_G$ with order n is transformed into a balanced state-space representation, whereby the *Gramian* matrices

$$\mathbf{W}_c = \mathbf{W}_o = \begin{bmatrix} \sigma_1 & & 0 \\ & \ddots & \\ 0 & & \sigma_n \end{bmatrix} \tag{4}$$

are diagonal and identical [22]. In [23] it was demonstrated that any complete controllable and observable system can be transformed into a balanced state-space representation. The diagonal elements σ_i of (4) are the *Hankel singular values* (HSV). The HSV can be determined by

$$\sigma_i = \sqrt{\lambda_i \mathbf{W}_c\mathbf{W}_o}. \tag{5}$$

Here, λ_i represents the eigenvalues of the system $\mathbf{\Sigma}_G$. The HSVs are invariant to transformations, so small HSVs correspond to badly controllable and observable states [22], [24]. The transformation matrix to the balanced state-space representation of (1) are determined in four steps:

1. In [25] it is shown that the controllability (2) and the observability matrix (3) are also solutions of the *Lyapunov* equation

$$\mathbf{A}\mathbf{W}_c + \mathbf{W}_c\mathbf{A}^T + \mathbf{B}\mathbf{B}^T = \mathbf{0} \tag{6}$$

$$\mathbf{A}^T\mathbf{W}_o + \mathbf{W}_o\mathbf{A} + \mathbf{C}^T\mathbf{C} = \mathbf{0} \tag{7}$$

2. After solving the *Lyapunov* equation, *Gramian* matrices are positive definite and can be decomposed into *Cholesky* factors: For the solution of the *Lyapunov* equation, one positive

definite solution for $\mathbf{W}_c$ and $\mathbf{W}_o$ exists for every fully controllable and observable system. The *Gramian* matrices can be decomposed into *Cholesky* factors

$$\mathbf{W}_c = \mathbf{S}^T\mathbf{S} \tag{8}$$

$$\mathbf{W}_o = \mathbf{R}^T\mathbf{R} \tag{9}$$

3. The product $\mathbf{SR}^T$ is subjected to a singular value decomposition (SVD),
$$\mathbf{SR}^T = \mathbf{U}\mathbf{\Sigma}\mathbf{V}^T, \tag{10}$$
which yields the orthogonal matrices $\mathbf{U}$ and $\mathbf{V}$, and the diagonal matrix $\mathbf{\Sigma}$ of HSVs.

4. The matrix $\mathbf{T}$ for transformation to balanced state-space representation is
$$\mathbf{T} = \mathbf{\Sigma}^{-\frac{1}{2}}\mathbf{V}^T\mathbf{R} \tag{11}$$
and its inverse is
$$\mathbf{T}^{-1} = \mathbf{S}^T\mathbf{U}\mathbf{\Sigma}^{-\frac{1}{2}} \tag{12}$$

In [26] it is presented that these matrices (11) and (12) transform a fully controllable and observable system in balanced state-space representation.

$$\mathbf{\Sigma}_{Bal} = \begin{cases} \dot{\mathbf{x}}_{Bal}(t) & = \mathbf{T}^{-1}\mathbf{A}\mathbf{T}\mathbf{x}_{Bal}(t) + \mathbf{T}^{-1}\mathbf{B}\mathbf{u}(t) \\ \mathbf{y}(t) & = \mathbf{C}\mathbf{T}\mathbf{x}_{Bal}(t) \end{cases} \tag{13}$$

The HSVs reveals that not all states are equally well controllable and observable. If only the states which contributes much to the energy transfer are used, then only the first r HSVs of $\mathbf{\Sigma}$ and the first r columns of $\mathbf{U}$ as well as $\mathbf{V}$ need to be considered. BT can also be interpreted as a *Petrov-Galerkin* projection into a reduced subspace [22]. For this projection, the transformation matrices (11) and (12) are modified.

$$\mathbf{W}_{BT}^T = \mathbf{\Sigma}^{-\frac{1}{2}}\mathbf{U}^T\mathbf{R}^T \tag{14}$$

$$\mathbf{V}_{BT} = \mathbf{S}\mathbf{V}\mathbf{\Sigma}^{-\frac{1}{2}} \tag{15}$$

Where $\mathbf{W}_{BT} \in \mathbb{R}^{r\times n}$ and $\mathbf{V}_{BT} \in \mathbb{R}^{r\times n}$ are the biorthogonal projection matrices. By the projection matrices only the first r columns are considered, thus the system order $r < n$ is reduced. This truncates only the HSVs with a small participation in the energy transfer, thus the input/output behavior is only marginally affected.

3.2 Krylov Subspaces

If numerically robust and efficient computations are required for model order reduction, the reduction must be performed by *Krylov* subspaces. Additionally, this method is used for very large systems. In [21] and [27] the *Krylov* subspace is generally defined. Let $\mathbf{M}$ be an arbitrary constant matrix and $\mathbf{v}$ a constant vector, called starting vector, then the *Krylov* subspace is defined as follows

$$\mathcal{K}_r(\mathbf{M}, \mathbf{v}) = \text{span}\left\{\mathbf{v}, \mathbf{M}\mathbf{v}, \ldots, \mathbf{M}^{r-1}\mathbf{v}\right\}. \tag{16}$$

Here, $\mathbf{M} \in \mathbb{R}^{n\times n}$, $\mathbf{v} \in \mathbb{R}^n$ are the arbitrary matrix and vector as well as $\mathcal{K}_r \subseteq \mathbb{R}^n$ is a r-dimensional subspace. For model order reduction with *Krylov* subspaces also transformation matrices $\mathbf{W}$ and $\mathbf{V}$ are used. The matrices are composed of column vectors, which are also the basis of the *Krylov* subspaces.
Referring to the system (1), the transformation matrix $\mathbf{V}$ as an arbitrary basis of the input subspace is determined by

$$\mathcal{K}_{r1}\left(\mathbf{A}^{-1}, \mathbf{A}^{-1}\mathbf{B}\right). \tag{17}$$

In addition, the transformation matrix $\mathbf{W}$ is derived as the basis of the output subspace

$$\mathcal{K}_{r2}\left(\mathbf{A}^{-T}, \mathbf{A}^{-T}\mathbf{C}^T\right). \tag{18}$$

For the reduction, $r_1 = r_2 = r$ must be required, whereby it must be ensured that both transformation matrices $\mathbf{V}$ and $\mathbf{W}$ have full rank r. The *Arnoldi* algorithm is used to determine the transformation matrices $\mathbf{V}$ and $\mathbf{W}$. Algorithm 1 shows the calculation of the transformation matrix $\mathbf{V}$ based on (17).

Algorithm 1: MIMO Arnoldi Algorithm [21]

Result: Transformation matrix $\mathbf{V}$
Set $\mathbf{v}_1 = \frac{\mathbf{b}_1}{\sqrt{\mathbf{b}_1^T\mathbf{b}_1}}$ and $\mathbf{b}_1$is the first starting vector;
for *i=2,3,...* **do**
  if $i \leq m_i$ **then**
    Next vector is the i-th starting vector
  else
    $\mathbf{r}_i = \mathbf{A}_1\mathbf{v}_{i-m_1}$
  end
  Set $\hat{\mathbf{v}}_i = \mathbf{r}_i$ ▷ Orthogonalization
  for $j = 1, \ldots, i-1$ **do**
    h = $\hat{\mathbf{v}}_i^T\mathbf{v}_j$
    $\hat{\mathbf{v}}_i = \hat{\mathbf{v}}_i - h\mathbf{v}_j$
  end
  if $\hat{\mathbf{v}}_i = 0$ **then**
    Reduce m_1 to $m_1 - 1$
    m_1 is nonzero calculate new $\mathbf{r}_i$
  else
    $\mathbf{v}_i = \frac{\hat{\mathbf{v}}_i}{\sqrt{\hat{\mathbf{v}}_i^T\hat{\mathbf{v}}_i}}$ ▷ i-th column of matrix $\mathbf{V}$
  end
end

The same algorithm and (18) is used to determine the matrix $\mathbf{W}$. The columns of the transformation matrices are spanning the corresponding *Krylov* input and output subspace (Fig. 1). If the model order reduction is performed based on (17) and (18) with matrices $\mathbf{V}$ and $\mathbf{W}$, the reduction procedure is called two-sided. Obviously, only one of the two matrices can be used. The remaining transformation matrix is chosen arbitrarily, but it must have full rank. The reduction procedure is then called one-sided.

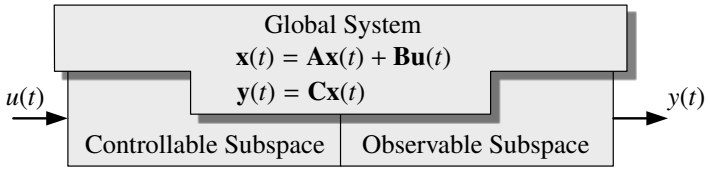

Figure 1: Controllability and observability are major properties and subspaces of dynamical control systems.

In *Krylov* subspace methods, the transfer functions of the original system and the reduced system are expanded at an arbitrary development point s_0, thus the first moments (coefficients of the *Taylor*

series) match. In this contribution, the development point $s_0 = 0$ is chosen, whereby the transfer function of (1) is described by a *Taylor* series around the developing point. The resulting i-th moment of the series around $s = 0$ is

$$\mathbf{m}_i = \mathbf{C}\mathbf{A}^{-(i+1)}\mathbf{B} \quad \text{with } i = 0, 1, \ldots \tag{19}$$

One-sided *Krylov* reduction methods reduce the original system to order r. In this case, the first r moments of the original system and the reduced system are matched. In the case of a two-sided method, the first $2r$ moments are matched.
The moment $\mathbf{m}_0$ corresponds to the stationary gain and is preserved in the *Krylov*-based reduction, so the reduced system is equally stationary accurate. Two-sided *Krylov* methods adjust two times as many moments for the same reduction size compared to one-sided methods. For this reason, two-sided methods are usually preferable. In addition, they offer the advantage that the reduction is independent of the representation of the original model. In contrast, regular state transformations and other equivalent transformations of the original model in one-sided reduction affect the transfer behavior in the reduced model [28]. Using a projection, the state vector $\mathbf{x}(t)$ is approximated by the reduced state vector $\mathbf{x}_r(t)$ as follows

$$\mathbf{x}(t) = \mathbf{V}\mathbf{x}_r(t). \tag{20}$$

Here, $\mathbf{V} \in \mathbb{R}^{n \times r}$ is the basis of the *Krylov* input subspace and is determined using algorithm 1. $\mathbf{x}_r \in \mathbb{R}^r$ is the state vector of the reduced system.
By applying the *Arnoldi* algorithm twice with different subspaces (17) or (18), the reduced system with the transformation matrices can be determined as follows.

$$\boldsymbol{\Sigma}_r = \begin{cases} \dot{\mathbf{x}}_r(t) &= \mathbf{W}^T\mathbf{A}\mathbf{V}\mathbf{x}_r(t) + \mathbf{W}^T\mathbf{B}\mathbf{u}(t) \\ \mathbf{y}_r(t) &= \mathbf{C}\mathbf{V}\mathbf{x}_r(t) \end{cases} \tag{21}$$

4 Model Distribution

Decentralization and distribution of global large-scale MIMO systems in a sensor and actuator network achieve order reduction, as individual nodes never need to estimate the entire global state vector. Often, this distribution occurs due to internal system relationships or due to physical system boundaries. In this section, we present two approaches for decentralization and distribution, both based on controllability and observability of the global system. One approach demonstrates decentralization into local SISO systems based on the inputs and outputs of the global MIMO system. The order of the local systems is reduced to further minimize the computational effort using model order reduction techniques. In a further approach, the local SISO systems will be generated based on the controllable and observable subspace. The order reduction resulting from decentralization and distribution is reflected in a reduced computational effort for each node. As a consequence, the required computational effort in a local node is lower than in a centralized system.

4.1 Input-/Output-based

By decentralizing the inputs and outputs of (1), m local systems with an equal number of sensors and actuators are derived. Thereby, the outputs of the global system (1) are decentralized without further considerations. To decentralize the inputs, the physical couplings must be considered. For this purpose, the decoupling method from [29] can be used. It is assumed that the inputs have a localized impact in large-scale systems. The inputs of the global system are decentralized into local systems as in [1]. Decentralization and distribution provides the ability to distribute the system in a network. In a fully connected network, minimal non-modeled couplings between the inputs are considered by communication.
Thus, m local SISO systems are derived from the global system $\boldsymbol{\Sigma}_G$. During decentralization, local systems with one input and one output each are expanded, but the system order n is not reduced (Fig. 2).

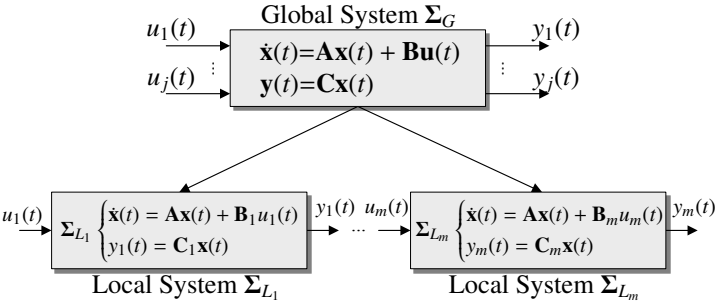

Figure 2: Decentralization of the global MIMO system with j inputs and outputs into m local systems.

The controllability and observability of the global system were not modified by decentralization. It can be demonstrated that by superposing the local *Gramian* controllability matrix $\mathbf{W}_{c_j}$ and the local observability matrix $\mathbf{W}_{o_j}$, the global controllability $\mathbf{W}_\mathbf{c}$ and observability $\mathbf{W}_o$ are recovered. The local systems are described as follows

$$\boldsymbol{\Sigma}_{L_j} = \begin{cases} \dot{\mathbf{x}}(t) = \mathbf{A}\mathbf{x}(t) + \mathbf{B}_j u_j(t) \\ y_j(t) = \mathbf{C}_j\mathbf{x}(t) \end{cases} \quad \text{with } j = 1, \ldots, m. \tag{22}$$

Where m describes the number of sensors and actuators. The index j at the input matrix $\mathbf{B}$ and at the output matrix $\mathbf{C}$ represents the corresponding column and row of the matrices. This defines local SISO systems $\boldsymbol{\Sigma}_{L_j}$ with one sensor and actuator each.
The local systems (22) are implemented on different decentralized CPUs, which communicate through a fieldbus system, and are used in a network.

4.1.1 Application in a Sensor and Actuator Network

Each local system $\boldsymbol{\Sigma}_{L_j}$ describes a network node where a Kalman Filter (KF) is used for state reconstruction [30]. The computational effort in the network nodes is further reduced if the KF algorithm is separated [1]. In the prediction step the system inputs are used and in the correction step, the observations are processed. Both filter steps communicate through a priori and a posteriori estimation $\hat{\mathbf{x}}(k)$ and covariance $\mathbf{P}(k)$. As the inputs are only applied in the prediction step, this part of the algorithm is implemented in the actuator nodes. In the sensor nodes, the observations are applied and only the correction step of the KF algorithm is executed. *Splitting the Kalman Filter* (SKF) provides the distribution of the local system (22) in a

network of sensor and actuator nodes. Each network node only has to calculate a part of the KF algorithm. By exchanging the a priori and a posteriori information, both filtering steps are connected in the network.
Therefore, the continuous-time dynamics of (22) must be transformed into discrete-time dynamics. To further minimize the computational effort, it is necessary to reduce the order of the discrete decentralized local systems. For this purpose, the model order reduction techniques presented in Section 3 are used.

$$\boldsymbol{\Sigma}_{L_{r,j}} = \begin{cases} \dot{\mathbf{x}}_r(t) = \mathbf{W}^T\mathbf{A}\mathbf{V}\mathbf{x}_r(t) + \mathbf{W}^T\mathbf{B}_j u_j(t) \\ y_{r,j}(t) = \mathbf{C}_j\mathbf{V}\mathbf{x}_r(t) \end{cases} \quad (23)$$
$$\text{with } j = 1, \ldots, m.$$

The decentralized reduced local systems (23) can also be used in the SKF. Depending on the reduction method, the projection matrices $\mathbf{V}$ and $\mathbf{W}$ are determined by Section 3.1 or Section 3.2. However, this turns the SKF into the *Splitted Reduced Kalman Filter* (SRKF)

Prediction

$$\hat{\mathbf{x}}_r^-(k) = \mathbf{A}_r(k-1)\hat{\mathbf{x}}_r^+(k-1) + \mathbf{B}_r(k-1)u_j(k-1) \quad (24)$$
$$\mathbf{P}_r^-(k) = \mathbf{A}_r(k-1)\mathbf{P}_r^+(k-1)\mathbf{A}_r^T(k-1) + \mathbf{Q}(k-1) \quad (25)$$

Correction

$$\mathbf{K}_r(k) = \mathbf{P}_r^-(k)\mathbf{C}_r^T(k)\left[\mathbf{C}_r(k)\mathbf{P}_r^-(k)\mathbf{C}_r^T(k) + \mathbf{R}(k)\right]^{-1} \quad (26)$$
$$\hat{\mathbf{x}}_r^+(k) = \hat{\mathbf{x}}_r^-(k) + \mathbf{K}_r(k)\left[y_j(k) - \mathbf{C}_r(k)\hat{\mathbf{x}}_r^-(k)\right] \quad (27)$$
$$\mathbf{P}_r^+(k) = \left[\mathbf{I} - \mathbf{K}_r(k)\mathbf{C}_r(k)\right]\mathbf{P}_r^-(k)\left[\mathbf{I} - \mathbf{K}_r(k)\mathbf{C}_r(k)\right]^T + \mathbf{K}_r(k)\mathbf{R}_r(k)\mathbf{K}_r^T(k). \quad (28)$$

The index r expresses that a reduced model $\boldsymbol{\Sigma}_{L_{r,j}}$ is used in the decentralized local nodes. The index j in (24) and (27) refers to the corresponding local system.
The distribution of prediction (24), (25) and correction (26),(27) as well as (28) of the local reduced systems (23) in sensor and actuator network nodes, is connected by the communication of the reduced a priori and a posteriori information. A fully connected network is formed if all m local reduced systems $\boldsymbol{\Sigma}_{L_{r,j}}$ communicate. This requires each sensor node to communicate its reduced a posteriori information to all local actuator nodes. This information is processed in the actuator nodes in a fusion step of the SRKF and used in determining the reduced a priori information. The SRKF becomes the *Decentralized Splitted Reduced Kalman Filter* (DSRKF) (Fig. 3). In [31], [32] and [33], a different notation is proposed for the a posteriori values, as the correction is performed using only the local observation.

$$\tilde{\mathbf{P}}_{r,j}^+(k) = \left[\left(\mathbf{P}_{r,j}^-(k)\right)^{-1}\mathbf{C}_{r,j}^T(k)\mathbf{R}_{r,j}^{-1}(k)\mathbf{C}_{r,j}(k)\right]^{-1} \quad (29)$$
$$\tilde{\mathbf{x}}_{r,j}^+ = \hat{\mathbf{x}}_{r,j}^-(k) + \mathbf{K}_{r,j}(k)\left[y_i(k) - \mathbf{C}_{r,j}(k)\hat{\mathbf{x}}_{r,j}^-\right] \quad (30)$$

For the interaction in the network between each local system, the information of the reduced covariance error and the reduced state error are used. These errors are determined in the corresponding local sensor nodes and transmitted to the actuator nodes. Thus, all actuator nodes receive the reduced error information of the sensor nodes. The reduced covariance error and reduced state error for the decentral reduced local system $\boldsymbol{\Sigma}_{L_{r,j}}$ is determined by

$$\mathbf{E}_{r,j}(k) = \left(\tilde{\mathbf{P}}_{r,j}^+(k)\right)^{-1} - \left(\mathbf{P}_{r,j}^-(k)\right)^{-1} \quad (31)$$
$$\mathbf{e}_{r,j}(k) = \left(\tilde{\mathbf{P}}_{r,j}^+(k)\right)^{-1}\tilde{\mathbf{x}}_{r,j}^+(k) - \left(\mathbf{P}_{r,j}^-(k)\right)^{-1}\hat{\mathbf{x}}_{r,j}^-(k) \quad (32)$$

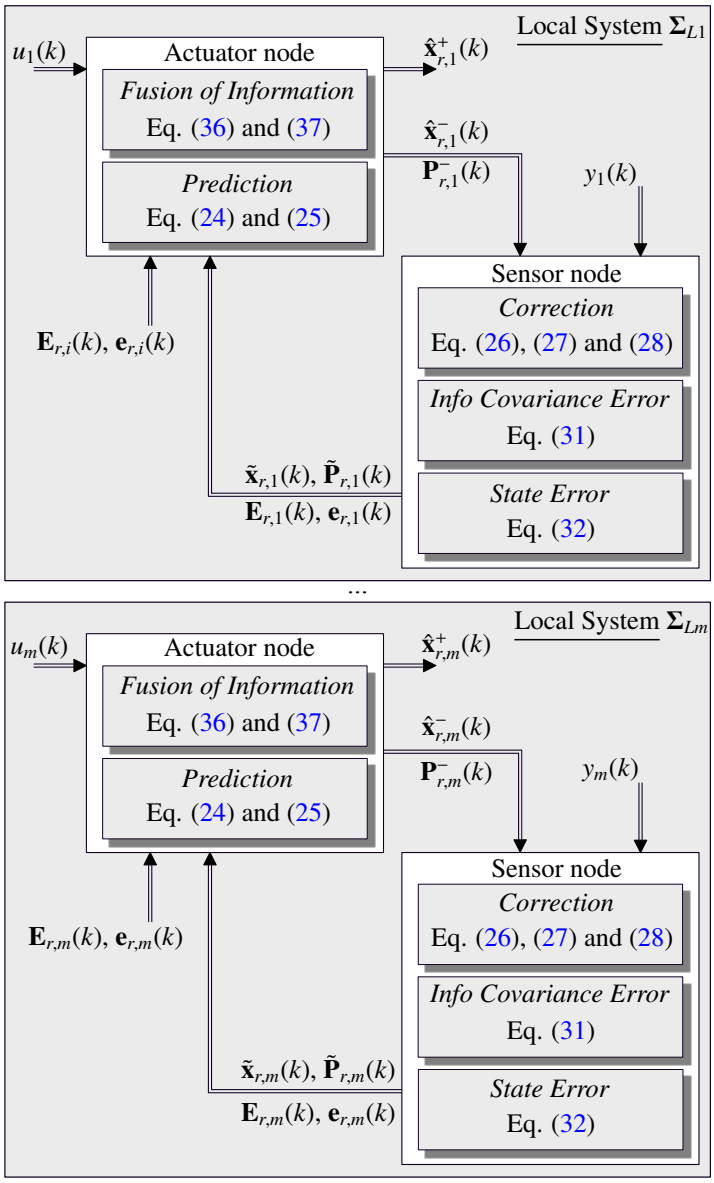

Figure 3: Splitting the KF algorithm on the actuator and sensor nodes. In all nodes reduced models are implemented (SRKF).

Each local system is projected by the matrices $\mathbf{V}$ and $\mathbf{W}$ into its own reduced subspace. In order to exchange information between the subspaces, the information, for instance the reduced error covariances $\mathbf{E}_{r,j}$ and the reduced state error vectors $\mathbf{e}_{r,j}$, must be transformed. The transformation from the transmitter subspace to the receiver subspace is performed in the receiving local actuator node. For the transformation matrices proposed in [31], we use the matrices already determined by the model order reduction in Section 3. These matrices are not symmetric, thus the *Moore-Penrose* inverse is used and specified by the superscript †. If the projection matrices

are determined by the method of Section 3.1, the transformation between the local subspaces are performed as follows

$$\mathbf{V}_{r,ji} = \mathbf{W}_{BT,i}\mathbf{W}^{\dagger}_{BT,j}. \tag{33}$$

For the state of clarity, the transformation using the methods form Section 3.2 is identical. As mentioned above, small couplings between inputs are considered in a fully networked sensor and actuator network. The error information of local system j are transformed by

$$\mathbf{E}_{r,ji}(k) = \left(\mathbf{V}^{T}_{r,ji}\right)^{\dagger}\mathbf{E}_{r,j}(k)\mathbf{V}^{\dagger}_{r,ji} \tag{34}$$

$$\mathbf{e}_{r,ji}(k) = \left(\mathbf{V}^{T}_{r,ji}\right)^{\dagger}\mathbf{e}_{r,j}(k) \tag{35}$$

into the subspace of the system i. By adding fusion equations

$$\left(\mathbf{P}^{+}_{r,j}(k)\right)^{-1} = \left(\mathbf{P}^{-}_{r,i}(k)\right)^{-1} + \sum_{j=1}^{m}\mathbf{E}_{r,ji}(k) \tag{36}$$

$$\hat{\mathbf{x}}^{+}_{r,j}(k) = \mathbf{P}^{+}_{r,j}(k)\left[\left(\mathbf{P}^{-}_{r,j}(k)\right)^{-1}\hat{\mathbf{x}}^{-}_{r,j}(k) + \sum_{j=1}^{m}\mathbf{e}_{r,ji}(k)\right] \tag{37}$$

to each actuator node, a fully interconnected network can be created.

Another approach uses for fusion all state estimates and covariance matrices of the decentralized reduced local systems (23). Here, the estimated a posteriori state vectors $\hat{\mathbf{x}}^{+}$ and a posteriori covariance matrix $\mathbf{P}^{+}$ are transmitted to a central fusion center. There is no direct communication between the locally reduced systems.
The fusion is performed in the center by the *Generalized Millman's Formula* (GMF). The GMF combines $\hat{\mathbf{x}}^{+}$ and $\mathbf{P}^{+}$ for multisensor systems [34], [35]. Assuming that m local systems estimate the vector $\mathbf{x} \in \mathbb{R}^{n}$ with $\hat{\mathbf{x}}$, these estimates are applied to determine the related local error covariance

$$\mathbf{P}_{i,j} = \mathrm{cov}\left(\tilde{\mathbf{x}}_i, \tilde{\mathbf{x}}_j\right) \tag{38}$$

$$\tilde{\mathbf{x}}_i = \mathbf{x} - \hat{\mathbf{x}}_i \tag{39}$$

$$i, j = 1, \ldots, m.$$

The objective of GMF is to find the optimal estimate of $\mathbf{x}$. [35] presents that this is achieved if the individual $\boldsymbol{\xi}_i$ are weighted and summed

$$\hat{\mathbf{x}} = \sum_{i=1}^{m}\boldsymbol{\xi}_i\hat{\mathbf{x}}_i. \tag{40}$$

In addition, it is shown that the weighting matrix $\boldsymbol{\xi}_i$ is found by minimizing the mean square error criterion, and that the minimization leads to linear equations

$$\sum_{i=1}^{m-1}\boldsymbol{\xi}_i\left(\mathbf{P}_{i,m-1} - \mathbf{P}_{i,m}\right) + \boldsymbol{\xi}_m\left(\mathbf{P}_{m,m-1} - \mathbf{P}_{m,m}\right) = \mathbf{0} \tag{41}$$

$$\sum_{i=1}^{m}\boldsymbol{\xi}_i = \mathbf{I}. \tag{42}$$

Therefore, the fusion of two local state vectors is proved to be equivalent to the well known *Bar-Shalom-Campo* formula. A more detailed background is given by [34] and [36].

Each decentralized reduced local system uses a KF for state reconstruction and communicates the estimated states and covariances to a fusion center (Fig. 4).

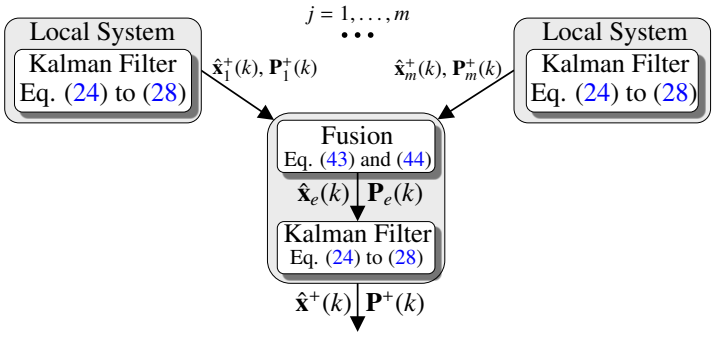

Figure 4: The state reconstruction in the local nodes is performed by a KF with a reduced model.

For independent estimates, GMF has a simple closed form

$$\hat{\mathbf{x}}_e = \mathbf{P}_e\sum_{j=1}^{m}\mathbf{P}^{+^{-1}}_{j}\hat{\mathbf{x}}^{+}_{j} \tag{43}$$

$$\mathbf{P}^{-1}_{e} = \sum_{j=1}^{m}\mathbf{P}^{+^{-1}}_{j}. \tag{44}$$

Due to decentralization and model order reduction, the decentralized reduced systems are transformed into different subspaces. The decentralized reduced estimates of the local systems have to be transformed into a common subspace for fusion. For this purpose, a reference system is derived from the global system (1). The reference system is described as follows

$$\boldsymbol{\Sigma}_{Ref} = \begin{cases}\dot{\mathbf{x}}_{Ref}(t) = \mathbf{W}^{T}\mathbf{A}\mathbf{V}\mathbf{x}_{Ref}(t) + \mathbf{W}^{T}\mathbf{B}\mathbf{u}(t)\\ \mathbf{y}_{Ref}(t) = \mathbf{C}\mathbf{V}\mathbf{x}_{Ref}(t).\end{cases} \tag{45}$$

In the model order reduction, the reference system is adapted to the order of the decentralized reduced local systems. In deriving the reference system, attention was given only to the stability and not to the quality of the approximation. At the fusion center, the transmitted information are considered as uncorrelated as they originate from different subspaces. After the received information is transformed into the reference system subspace, it is fused (43) and (44). Based on the reference model, a KF is implemented in the fusion center. This interprets the fusion $\hat{\mathbf{x}}_e$ as an observation and this improves the fusion result.

4.2 Subspace based

Using the input subspace $\mathcal{K}_{r1}$ and the output subspace $\mathcal{K}_{r2}$ separately as well as algorithm 1, the global system $\boldsymbol{\Sigma}_G$ is reduced and distributed. Here, a one-sided method is chosen, thus the controllable subspace is determined only by using the transformation matrix $\mathbf{V}$ [37]. For reduction, $\mathbf{W}^{*} = \mathbf{V}$ is chosen so that $\mathbf{W}^{*}$ has maximum rank. Thus, a reduced order model is found,

$$\boldsymbol{\Sigma}_{r,C} = \begin{cases}\dot{\mathbf{x}}_{r,C}(t) = \mathbf{W}^{*^{T}}\mathbf{A}\mathbf{V}\mathbf{x}_{r,C}(t) + \mathbf{W}^{*^{T}}\mathbf{B}\mathbf{u}(t)\\ y_{r,C}(t) = \mathbf{C}\mathbf{V}\mathbf{x}_{r,C}(t)\end{cases} \tag{46}$$

which describes the controllable subspace. The application of algorithm 1 is repeated for the *Krylov* output space. Thus the matrix $\mathbf{W}$ is determined. For this reduction $\mathbf{V}^* = \mathbf{W}$ is chosen, thus also here the second matrix has maximal rank. The reduced model

$$\mathbf{\Sigma}_{r,O} = \begin{cases} \dot{\mathbf{x}}_{r,O}(t) = & \mathbf{W}^T\mathbf{A}\mathbf{V}^*\mathbf{x}_{r,O}(t) + \mathbf{W}^T\mathbf{B}\mathbf{u}(t) \\ y_{r,O}(t) = & \mathbf{C}\mathbf{V}^*\mathbf{x}_{r,O}(t) \end{cases} \tag{47}$$

describes the observable subspace. Using (46) and (47), the global system is approximated over the controllable and observable subspace. A reduced system is obtained by the system matrix $\mathbf{A}_{r,C}$ and the input matrix $\mathbf{B}_{r,C}$, of the controllable subspace, and by the output matrix $\mathbf{C}_{r,O}$, of the observable subspace

$$\begin{aligned} \mathbf{A}_{r,C} &= \mathbf{W}^{*^T}\mathbf{A}\mathbf{V} \\ \mathbf{B}_{r,C} &= \mathbf{W}^{*^T}\mathbf{B} \\ \mathbf{C}_{r,O} &= \mathbf{C}\mathbf{V}^* \end{aligned} \tag{48}$$

4.2.1 Application in a Sensor and Actuator Network

A KF is also used for state reconstruction in the subspace-based distribution. After the transformation from continuous to discrete time, the system is used in KF. By distributing into the controllable and observable subspace, this method can also be used in a network. Appling the SKF presented in Section 4.1, the computational effort in the decentralized CPUs is reduced because each network node only needs to determine a part of the KF algorithm. By using the reduced model (48), the SKF becomes the SRKF as described above (Fig. 5). However, information from both subspaces is necessary for a complete approximation. By communicating a priori estimates $\hat{\mathbf{x}}_r^-$ and covariances matrices $\mathbf{P}_r^-$ as well as a posteriori state vectors $\hat{\mathbf{x}}_r^+$ and covariance matrices $\mathbf{P}_r^+$, a interconnected network is created. To achieve this objective, the orthonormal bases from Section 3 are used. For communication between the reduced controllable and reduced observable subspace, the transmitted data must be transformed into the target subspace.

$$\mathbf{T}_c^o = \mathbf{W}\mathbf{V}^\dagger. \tag{49}$$

The index of $\mathbf{T}$ describes the starting point and the superscript indicates the target point of transformation. Due to the lack of symmetry of the orthonormal bases, the *Moore-Penrose* inverse is here also necessary. As in the previous section, we assume that no packets are lost and that both nodes operate with the same sampling time.

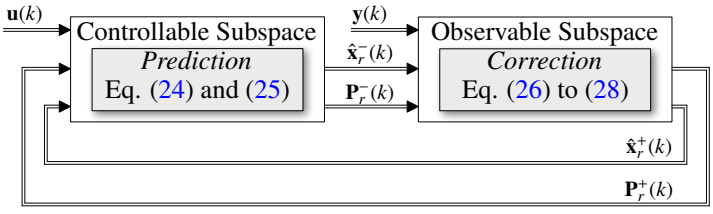

Figure 5: Distribution of the global linear system on the basis of controllability and observability

5 Experimental Evaluation and Results

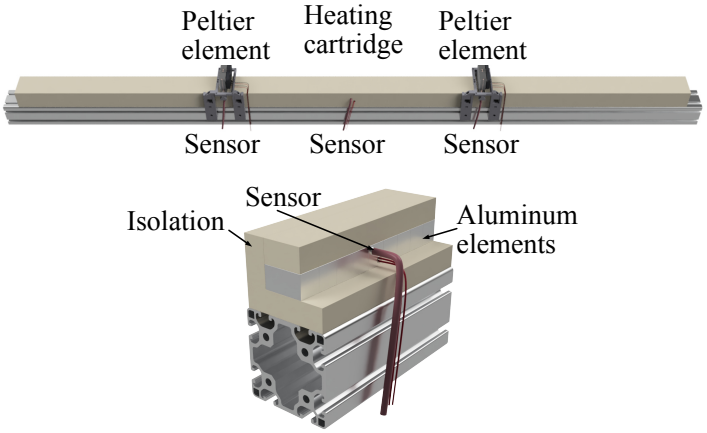

Figure 6: Visualization of a large scaled system. The discrete aluminum rod (l = 2m) is coold on two elements and heated on another.

We have experimentally evaluated the presented methods on an example published in [1] and [2]. Here we analyzed the temperature distribution in a discrete heating rod.
The aluminum rod consists of $n = 100$ elements. The contact surfaces of the aluminum elements have been thermally insulated, minimizing the influence of ambient temperature. The heat conducting rod is cooled by two Peltier elements and heated by a heating cartridge. In addition, the temperature distribution in the aluminum elements is monitored by three PT1000 sensors. The global system of the thermal rod can be described as follows

$$\begin{aligned} &\mathbf{x}_q(k+1) = p_1\mathbf{x}_{q-1}(k) + (2p_2 + p_3)\,\mathbf{x}_q(k) + p_1\mathbf{x}_{q+1}(k) + \mathbf{w}_q \\ &\quad \text{with } q = 1, \ldots, n \end{aligned} \tag{50}$$

Where the index q is the number of the aluminum element. The parameters p_1 and p_2 represents the element geometry, heat capacity as well as the heat transfer coefficient. The elements $q = 1$ and $q = 100$ have only one neighboring element, thus the heat emitting surface is larger than for the other elements. For this purpose, $2p_2 + p_3 \Longrightarrow p_2 + p_4$ is chosen for the elements at the end of the rod. The parameter p_3 and p_4 are the product of the heat transfer coefficient and the heat emitting surface. $\mathbf{w}_q$ is a normally distributed noise, describing uncertainties in the process.
The system order is $n = 100$ and is determined by the elements of the rod. Thus the system matrix is $\mathbf{A} \in \mathbb{R}^{100\times100}$. Here, the system has three actuators and three sensors, so the input matrix is $\mathbf{B} \in \mathbb{R}^{n\times3}$ and the output matrix is $\mathbf{C} \in \mathbb{R}^{3\times n}$. The sensors and actuators are placed in the same aluminum element (25, 50, 75). So the input matrix $\mathbf{B}$ is a zero matrix and is only assigned at the positions $B_{25,1} = -1$, $B_{50,2} = 1$ and $B_{75,3} = -1$ and the output matrix $\mathbf{C}$ is also a zero matrix and is also only assigned at the elements $C_{1,25} = 1$, $C_{2,50} = 1$ and $C_{3,75} = 1$. The global system can be described as follows

$$\mathbf{\Sigma}_{Rod} = \begin{cases} \mathbf{x}(k+1) & = \mathbf{A}\mathbf{x}(k) + \mathbf{B}\mathbf{u}(k) \\ \mathbf{y}(k) & = \mathbf{C}\mathbf{x}(k) \end{cases} \tag{51}$$

5.1 Parameter Identification

The unknown parameters of the global system are determined by identification in two steps. In the first step, the parameter p_2 is identified by the homogeneous solution of the state space equation. For this purpose, a single aluminum element is used to determine the main diagonal of the system matrix $\mathbf{A}$ (Fig. 7).

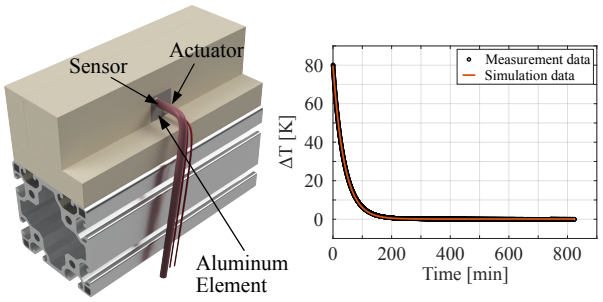

Figure 7: Parameter identification of the main diagonal elements.

This element is equipped with the heating cartridge and a sensor, and the isolation minimizes the influence of the ambient temperature. The heating cartridge will heat the temperature in the element to T=80°C. When the temperature reaches T=80°C, the power of the heating cartridge is switched off and the cooling curve is recorded until the steady state is reached (Fig. 7).
For the identification procedure, a first-order linear model and a *Nelder-Meat* simplex are used in combination with an ode4 solver. The least squares method is used to evaluate the output error and the *Nelder-Meat* simplex minimizes the cost function

$$\min J = \min_{\hat{\mathbf{p}} \in \mathbb{R}^n} \sum_{k=1}^{N} \left[y(k) - \hat{y}(k) \right]^2 . \tag{52}$$

so that the behavior of the model is fitted to the real system.
In the second identification step, the secondary diagonal elements of the system matrix $\mathbf{A}$ are determined. For this purpose, seven aluminum elements are used to reduce the discretization error (Fig. 8). The elements at the edge are equipped with a sensor. The middle element is also equipped together with the heating cartridge. With constant heating power, the elements are heated until the steady state is reached. A seventh order model is used for identification and least squares is used to evaluate the output error. The cost function (52) is also minimized using a *Nelder-Meat* simplex in combination with an ode4 solver.

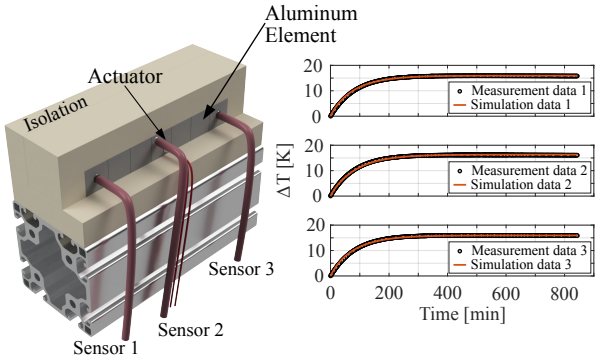

Figure 8: Parameter identification of the secondary diagonal elements.

These identified parameters can be used in the global system

$$\begin{aligned} p_1 &= 4.2812 \cdot 10^{-4} \tfrac{1}{s} \quad p_2 = 4.3048 \cdot 10^{-4} \tfrac{1}{s} \\ p_3 &= 1.8067 \cdot 10^{-4} \tfrac{1}{s} \quad p_4 = 2.2561 \cdot 10^{-4} \tfrac{1}{s} \end{aligned} \tag{53}$$

5.2 Input-/Output-based Distribution

The starting point is a global MIMO system (51) with three actuators and three sensors and a system order of $n = 100$. Considering the controllability and observability of the global system $\mathbf{\Sigma}_{Rod}$, $m = 3$ local systems are decentralized with one actuator and sensor each.

$$\begin{aligned} \mathbf{\Sigma}_{L_1} &\begin{cases} \mathbf{x}(k+1) &= \mathbf{A}\mathbf{x}(k) + \mathbf{B}_j u_1(k) \\ y_1(k) &= \mathbf{C}_j \mathbf{x}(k) \end{cases} \\ \mathbf{\Sigma}_{L_2} &\begin{cases} \mathbf{x}(k+1) &= \mathbf{A}\mathbf{x}(k) + \mathbf{B}_j u_2(k) \\ y_2(k) &= \mathbf{C}_j \mathbf{x}(k) \end{cases} \\ \mathbf{\Sigma}_{L_3} &\begin{cases} \mathbf{x}(k+1) &= \mathbf{A}\mathbf{x}(k) + \mathbf{B}_j u_3(k) \\ y_3(k) &= \mathbf{C}_j \mathbf{x}(k) \end{cases} \end{aligned} \tag{54}$$

The index $j = 1, \ldots, m$ is the corresponding column of the global input matrix $\mathbf{B}$ and the corresponding row of the global output matrix $\mathbf{C}$. The system order is not changed, thus the decentralized local systems are also $n = 100$.
Analyzing the HSVs reveals, that order $r = 6$ is achieved (Fig. 9). The order $r = 6$ allows 93.1217% of the energy to be transported from the inputs to the outputs. Based on the analysis, states with a small impact on the energy transport are cut off and the system order is reduced to $r = 6$. The model order reduction reduces the computational effort in the network node.
The decentralized reduced local systems are extended by the DSRKF presented in Section 4, providing communication between the network nodes and generating a fully interconnected network.

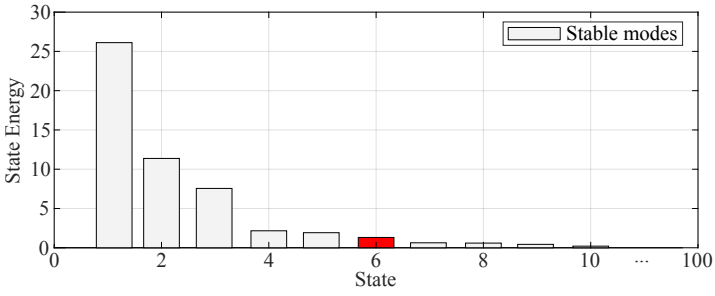

Figure 9: Visualization of the state energy of the global system (50). It is illustrated that the energy decreases strongly with the sixth state. Red bar illustrates the cut-off system order r.

Fig. 3 illustrates that the sensor nodes communicate with the actuator nodes by means of the exchanged a priori and a posteriori information. The full interconnectivity of the network is achieved by exchanging the error information (31) and (32) between the decentralized reduced local systems. This error information is already used in the prediction, so a fusion step has to be added to the algorithm in the actuator nodes (36) and (37).
By using BT, the physical interpretability of the states is lost, thus, for an analysis, the reduced state vector has to be transformed back

to the global state space. For this transformation of the a priori and a posteriori information, the following equations are used

$$\hat{\mathbf{x}}^{-}(k) = \mathbf{V}_{BT}\hat{\mathbf{x}}_r^{-}(k) \quad (55)$$

$$\hat{\mathbf{x}}^{+}(k) = \mathbf{V}_{BT}\hat{\mathbf{x}}_r^{+}(k). \quad (56)$$

The transformation matrix $\mathbf{V}_{BT}$ is determined by (15).

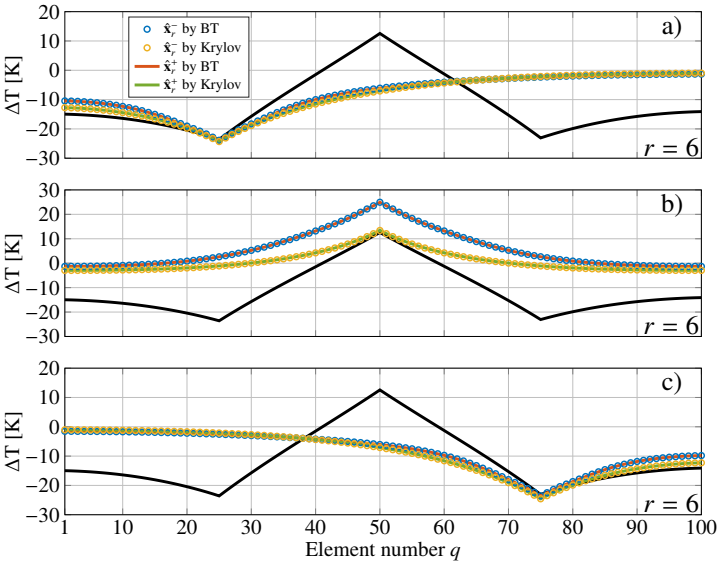

Figure 10: A priori and a posteriori estimated state vectors of local systems reduced by BT (3.1) and *Krylov* (3.2).

Fig. 10 demonstrates that the local SISO systems satisfy the characteristics of the decentralized inputs and outputs of the global system. Hereby, the global temperature profile (50) is shown in black and the a priori estimation $\hat{\mathbf{x}}_r^-$ after the reduction by BT/Krylov of the acutator node, is represented in blue/orange dots as well as the a posteriori estimation $\hat{\mathbf{x}}_r^+$ in red/green lines of the sensor node. In a), b) and c) the estimates based on measurement data of the reduced local systems after the transformation into the original SISO subspace ($n = 100$) are illustrated.

In addition, it is presented that despite the lost interpretability of the states due to the model order reduction as well as the distribution of the estimation algorithms in a interconnected network, the controllability and observability of the global system is approximated by the decentralized reduced local systems.

The input-/output-based distribution from Section 4.1 is also analyzed using the model order reduction method from Section 3.2, the two-sided *Krylov* method. Although the model order reduction with *Krylov* subspace methods is numerically efficient, the stability of the reduced system is not guaranteed. Therefore, the stability must be checked in the preprocessing. Based on the stability analysis of the decentralized local systems, the reduction with *Krylov* subspace methods also yields stable reduced systems for $r = 6$.

This approach also illustrates that the characteristics of the sensors and actuators of the global system is approximated by the decentralized reduced local systems (Fig. 10). Furthermore, the described method for decentralization is used independently of the two model order reduction methods presented.

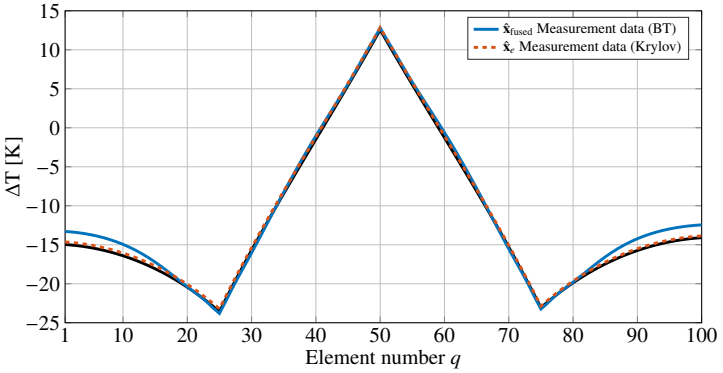

Figure 11: By BT and *Krylov* reduced estimated a posteriori state vectors with (57) and (58) as well as (43) and (44) fused in a center.

If the a posteriori information of the decentralized SISO system reduced by BT is provided to a fusion center, the original $\mathbf{\Sigma}_{Rod}$ is reconstructed. For this purpose, the reduced a posteriori state vector $\hat{\mathbf{x}}_r^+$ and the a posteriori covariance matrix $\mathbf{P}_r^+$ of the $m = 3$ local systems must be transformed into the local model space and fused by

$$\mathbf{P}_{\text{fused}}(k) = \left(\sum_{j=1}^{m} \mathbf{P}_j^{-}(k)\right)^{-1} \quad (57)$$

$$\hat{\mathbf{x}}_{\text{fused}}(k) = \mathbf{P}_{\text{fused}}(k) \sum_{j=1}^{m} \left(\mathbf{P}_j^{-}(k)\right)^{-1} \hat{\mathbf{x}}_j^{-}(k). \quad (58)$$

For the reconstruction of the global system $\mathbf{\Sigma}_{Rod}$ based on the second method of model order reduction, a reference system (45) is derived. The stability and the equivalent system order ($r = 6$) as for the $m = 3$ decentralized reduced local systems have to be considered. Before the decentrally reduced local a posteriori information is used for fusion, a transformation into a known common subspace is required. For this purpose, the decentrally reduced local a posteriori information is transformed into the subspace of the reference system. For further improvement a KF is used , which achieves an increase of the reconstruction $\hat{\mathbf{x}}_e$ of the global system $\mathbf{\Sigma}_{Rod}$ with the reference system and the fusion result as observation.

In contrast to Fig. 10, a fusion center is used in Fig. 11. The global temperature profile according to (50) with the identified parameters is shown in black. The decentralized local a posteriori states, reduced by BT, are shown in blue after fusion by (57) and (58) and transformation ($n = 100$), while the dashed red line represents the decentralized local a posteriori states, reduced by the method of *Krylov* subspaces.

Here, the global system is fully reconstructed despite decentralization and distribution in a sensor and actuator network. Additionally, it is demonstrated that the presented methods are independent of the model order reduction techniques. The decentralization and distribution would be implemented for both model order reduction methods (BT and *Krylov* subspace methods) on BECKHOFF® CX6030 CPUs (Intel® Core™ i7 7700, 3.6GHz) and tested in real-time approaches.

The deviation of the approximation by Balanced Truncation in the ranges $q = 1, \ldots, 25$ and $q = 75, \ldots, 100$ is accounted for by the reduction of the system order. With an increase of the order this

deviation is minimized.
The communication effort is directly dependent on the system order. For clarification, the system (51) was manually distributed into several local systems. The order of the local systems is $n_1 = 35$, $n_2 = 20$ and $n_3 = 45$. If these decentralized local systems are used to reconstruct the global system without further reduction, all local systems must communicate with each other. The communication overhead is significantly higher than in the presented methods with the reduction methods BT and Krylov subspace methods (Fig. 12). Thus, the presented methods reduce not only the computational effort in the network nodes but also the communication effort.

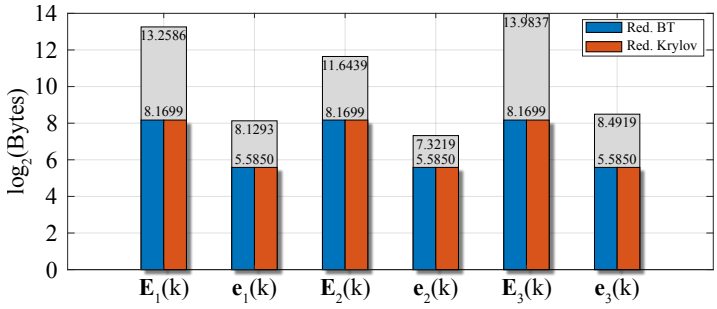

Figure 12: The information of the covariance error and state error sent by manual splitting is illustrated as grey and for DSRKF as blue/red bars (logarithmic scale).

5.3 Subspace based Distribution

Without decentralization, the controllable and observable *Krylov* subspace is derived directly from the global system (51). For this purpose, the repeating of the one-sided method for the *Krylov* input space and the *Krylov* output space is proposed in Secton 4.2. The resulting reduced systems (46) and (47) describe the corresponding subspace. However, to approximate the global input and output behavior, information from both subspaces are required. For state reconstruction in the subspace-based distribution, a SRKF from Section 4.1.1 is used with the matrices (48). By separating the controllable and observable subspace, this distribution can also be used in a network.

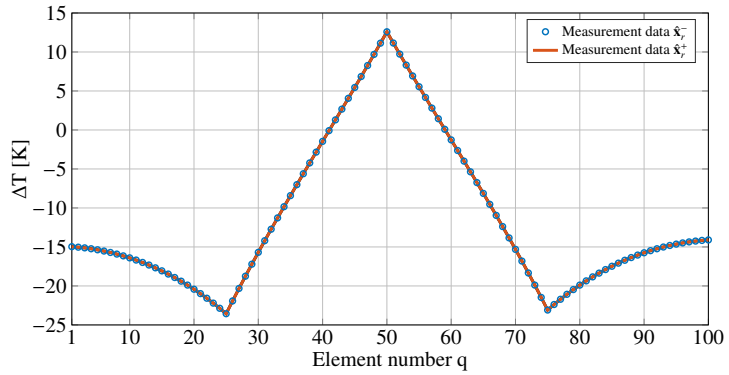

Figure 13: Reconstructing the global temperature profile by communicating the reduced a priori and a posteriori states according to the distribution of the controllability and observability subspace in a sensor and actuator network.

The network nodes transmit the a priori and a posteriori state vectors and covariance matrix, so no fusion steps are required in the algorithms. However, the information received in the subspaces must be transformed (49). Even when *Krylov* methods are used without decentralization, the stability of the reduced system must be analyzed in preprocessing. Fig. 13 demonstrates that also in this case, for $r = 6$, a stable reduced system is present and used.
In contrast to the previous method (Section 5.2), only two network nodes are used here. One network node uses the complete information of the inputs for the prediction of the KF and another network node uses all observations for the correction step of the reconstruction of the global system behaviour.
The minimization of the communication effort is not depicted, as in this distribution only the prediction and correction is separated (SKF). Using the model order reduction (SRKF), there will be a lower communication effort with each reduced system order, since a manual division cannot be performed when separating prediction and correction.

6 Conclusion

In this paper, we have shown the decentralization into SISO systems of large-scale MIMO systems. We have also presented an input/output-based and a subspace-based method for system distribution. Both methods focus on the controllability and observability of the global MIMO system. By applying the two best known model order reduction methods (balanced truncation and *Krylov* subspaces), the system order of the generated decentralized local systems was reduced. By adding a DSRKF to the decentralized reduced local systems, they are used in a network. The reduced decentralized local systems describe different subspaces, so it is assumed that the reduced information is communicated in the network. It could be demonstrated that the transformation matrices resulting from the two presented methods can also be used for communication. Their distribution in a connected network in combination with the model order reduction methods significantly reduced the computational and communication effort required for state estimation of the large-scale linear MIMO system.
In the future, the presented methods will also be applied for decentralized distributed control. In future practical approaches, the methods will also be used in tracking dynamic targets and applied to the decentralized distributed multi-laser tracker system (DDMLTS) [38].

Conflict of Interest The authors declare no conflict of interest.

References

[1] F. Friedrich, J. Mayer, C. Ament, "Reduced and Distributed Estimation in Sensor and Actuator Networks - Automated Design Based on Controllability and Observability," in 2021 IEEE Conference on Control Technology and Applications (CCTA), 666–672, IEEE, San Diego, CA, USA, 2021, doi: 10.1109/CCTA48906.2021.9659129.

[2] F. Friedrich, J. Mayer, C. Ament, "Model-Order Reduction and System Distribution Using Krylov Subspaces - An Approach for Efficient State Estimation in Sensor and Actuator Networks," accepted for CCTA2022.

[3] R. D'Andrea, G. Dullerud, "Distributed control design for spatially interconnected systems," IEEE Transactions on Automatic Control, **48**(9), 1478–1495, 2003, doi:10.1109/TAC.2003.816954.

[4] B. Bamieh, F. Paganini, M. Dahleh, "Distributed control of spatially invariant systems," IEEE Transactions on Automatic Control, **47**(7), 1091–1107, 2002, doi:10.1109/TAC.2002.800646.

[5] L. A. Montestruque, P. J. Antsaklis, "Model-Based Networked Control Systems - Stability," 58.

[6] J. Lunze, Regelungstechnik 2, Springer Berlin Heidelberg, Berlin, Heidelberg, 2016, doi:10.1007/978-3-662-52676-7.

[7] U. A. Khan, J. M. F. Moura, "Model Distribution for Distributed Kalman Filters: A Graph Theoretic Approach," in 2007 Conference Record of the Forty-First Asilomar Conference on Signals, Systems and Computers, 611–615, IEEE, Pacific Grove, CA, USA, 2007, doi:10.1109/ACSSC.2007.4487286, iSSN: 1058-6393.

[8] B. Noack, J. Sijs, U. D. Hanebeck, "Fusion Strategies for Unequal State Vectors in Distributed Kalman Filtering," IFAC Proceedings Volumes, **47**(3), 3262–3267, 2014, doi:10.3182/20140824-6-ZA-1003.02491.

[9] J. Lunze, Control Theory of Digitally Networked Dynamic Systems, Springer International Publishing, Heidelberg, 2014, doi:10.1007/978-3-319-01131-8.

[10] A. C. Antoulas, Approximation of large-scale dynamical systems, Advances in design and control, Society for Industrial and Applied Mathematics, Philadelphia, 2005, doi:10.1137/978-0-89871-871-3.

[11] P. Benner, A. Schneider, "Balanced Truncation Model Order Reduction for LTI Systems with many Inputs or Outputs," in Proceedings of the 19th International Symposium on Mathematical Theory of Networks and Systems – MTNS 2010, 4, MTNS, Budapest, Hungary, 2010.

[12] E. F. Yetkin, H. Dag, "Parallel implementation of iterative rational Krylov methods for model order reduction," in 2009 Fifth International Conference on Soft Computing, Computing with Words and Perceptions in System Analysis, Decision and Control, 1–4, IEEE, Famagusta, 2009, doi:10.1109/ICSCCW.2009.5379421.

[13] Z. Bai, "Krylov subspace techniques for reduced-order modeling of large-scale dynamical systems," Applied Numerical Mathematics, **43**(1-2), 9–44, 2002, doi:10.1016/S0168-9274(02)00116-2.

[14] P. Feldmann, R. Freund, "Efficient linear circuit analysis by Pade approximation via the Lanczos process," IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems, **14**(5), 639–649, 1995, doi:10.1109/43.384428.

[15] K. Gallivan, E. Grimme, P. Van Dooren, "Asymptotic waveform evaluation via a Lanczos method," Applied Mathematics Letters, **7**(5), 75–80, 1994, doi:10.1016/0893-9659(94)90077-9.

[16] G. A. Baker, P. R. Graves-Morris, Padé approximants, number v. 59 in Encyclopedia of mathematics and its applications, Cambridge University Press, Cambridge [England] ; New York, 2nd edition, 1996.

[17] D. Rafiq, M. A. Bazaz, "Model Order Reduction via Moment-Matching: A State of the Art Review," Archives of Computational Methods in Engineering, **29**(3), 1463–1483, 2022, doi:10.1007/s11831-021-09618-2.

[18] L. Feng, "Review of model order reduction methods for numerical simulation of nonlinear circuits," Applied Mathematics and Computation, **167**(1), 576–591, 2005, doi:10.1016/j.amc.2003.10.066.

[19] H. K. F. Panzer, T. Wolf, B. Lohmann, "$\mathcal{H}_2$ and $\mathcal{H}_\infty$ error bounds for model order reduction of second order systems by Krylov subspace methods," in 2013 European Control Conference (ECC), 4484–4489, IEEE, Zurich, 2013, doi:10.23919/ECC.2013.6669657.

[20] S. Gugercin, Projection Methods for Model Reduction of Large-Scale Dynamical Systems, Dissertation, Rice University, Rice, 2003.

[21] B. Salimbahrami, B. Lohmann, T. Bechtold, "Two-Sided Arnoldi in Order Reduction of Large Scale MIMO Systems," 9.

[22] B. Moore, "Principal component analysis in linear systems: Controllability, observability, and model reduction," IEEE Transactions on Automatic Control, **26**(1), 17–32, 1981, doi:10.1109/TAC.1981.1102568.

[23] S. Gugercin, A. C. Antoulas, "A Survey of Model Reduction by Balanced Truncation and Some New Results," International Journal of Control, **77**(8), 748–766, 2004, doi:10.1080/00207170410001713448.

[24] D. Hinrichsen, H.-W. Philippsen, "Modellreduktion mit Hilfe balancierter Realisierungen / Model reduction using balanced realizations," at - Automatisierungstechnik, **38**(1-12), 460–466, 1990, doi:10.1524/auto.1990.38.112.460, publisher: De Gruyter (O).

[25] S. L. Brunton, J. N. Kutz, Data-Driven Science and Engineering, Cambridge University Press, 1st edition, 2019.

[26] A. Laub, M. Heath, C. Paige, R. Ward, "Computation of system balancing transformations and other applications of simultaneous diagonalization algorithms," IEEE Transactions on Automatic Control, **32**(2), 115–122, 1987, doi:10.1109/TAC.1987.1104549, conference Name: IEEE Transactions on Automatic Control.

[27] I. Jaimoukha, "A general minimal residual Krylov subspace method for large-scale model reduction," IEEE Transactions on Automatic Control, **42**(10), 1422–1427, 1997, doi:10.1109/9.633831.

[28] B. Lohmann, B. Salimbahrami, "Introduction to Krylov subspace methods in model order reduction," Methods Applications in Automation, 1–13, 2003.

[29] P. Falb, W. Wolovich, "Decoupling in the design and synthesis of multivariable control systems," IEEE Transactions on Automatic Control, **12**(6), 651–659, 1967, doi:10.1109/TAC.1967.1098737.

[30] D. Simon, Optimal state estimation: Kalman, H [infinity] and nonlinear approaches, Wiley-Interscience, Hoboken, N.J, 2006, doi:10.1002/0470045345.

[31] A. G. Mutambara, Decentralized Estimation and Control for Multisensor Systems, Routledge, 1st edition, 2019, doi:10.1201/9781315140803.

[32] P. Hilgers, C. Ament, "Distributed and decentralised estimation of non-linear systems," in 2010 IEEE International Conference on Control Applications, 328–333, 2010, doi:10.1109/CCA.2010.5611300, iSSN: 1085-1992.

[33] B. S. Rao, H. F. Durrant-Whyte, "Fully decentralised algorithm for multisensor Kalman filtering," IEE Proceedings D (Control Theory and Applications), **138**(5), 413–420, 1991, doi:10.1049/ip-d.1991.0057, publisher: IET Digital Library.

[34] V. Shin, Y. Lee, T.-S. Choi, "Generalized Millman's formula and its application for estimation problems," Signal Processing, **86**(2), 257–266, 2006, doi:10.1016/j.sigpro.2005.05.015.

[35] H. B. Mitchell, Multi-sensor data fusion: an introduction, Springer, Berlin ; New York, 2007, doi:0.1007/978-3-540-71559-7.

[36] J. Ajgl, "Millman's Formula in Data Fusion," in The 10th International PhD Workshop Young Generation Viewpoint, 6.

[37] B. Salimbahrami, R. Eid, B. Lohmann, "On the choice of an optimal interpolation point in Krylov-based order reduction," in 2008 47th IEEE Conference on Decision and Control, 4209–4214, IEEE, Cancun, 2008, doi:10.1109/CDC.2008.4739219.

[38] F. Friedrich, J. Brandl, C. Ament, "Absolute Distance Measurement by a Decentralized and Distributed Multi-Lasertracker-System," in 2022 IEEE/ASME International Conference on Advanced Intelligent Mechatronics (AIM), 1249–1255, IEEE, Sapporo, Japan, 2022, doi:10.1109/AIM52237.2022.9863275.