Simultaneous State and Parameter Estimation of Distributed-Parameter Physical Systems based on Sliced Gaussian Mixture Filter

Felix Sawo, Vesa Klumpp, and Uwe D. Hanebeck
Intelligent Sensor-Actuator-Systems Laboratory
Institute of Computer Science and Engineering
Universität Karlsruhe (TH), Germany
Email: {sawo, klumpp}@ira.uka.de, uwe.hanebeck@ieee.org

***Abstract*—This paper presents a method for the simultaneous state and parameter estimation of finite-dimensional models of distributed systems monitored by a sensor network. In the first step, the distributed system is spatially and temporally decomposed leading to a linear finite-dimensional model in state space form. The main challenge is that the simultaneous state and parameter estimation of such systems leads to a high-dimensional nonlinear problem. Thanks to the *linear* substructure contained in the resulting finite-dimensional model, the development of an overall more efficient estimation process is possible. Therefore, in the second step, we propose the application of a novel density representation – *sliced Gaussian mixture density* – in order to decompose the estimation problem into a (conditionally) linear and a nonlinear problem. The systematic approximation procedure minimizing a certain distance measure allows the derivation of (close to) optimal and deterministic results. The proposed estimation process provides novel prospects in sensor network applications. The performance is demonstrated by means of simulation results.**

Keywords: Distributed systems, simultaneous state and parameter estimation, sensor networks, nonlinear estimation.

I. Introduction

Recent developments and miniaturization of sensor nodes make it possible to use a wireless sensor network for monitoring natural large-area phenomena. In such scenarios, the individual sensor nodes are densely deployed, either inside the phenomenon or close to it. By the distribution of local information through the sensor network the phenomenon to be observed can be coöperatively reconstructed in an intelligent and autonomous manner [1], [2].

That means, the sensor network can be regarded as a huge information field collecting data from its surrounding and then providing useful information both to mobile agents and to humans. Based on the extended perception provided by the sensor network they would be able to accomplish certain tasks more efficiently or could be warned in dangerous situations, such as avalanches, forest fires or seismic sea waves. Examples of such distributed physical quantities could be: temperature distributions, chemical concentrations, fluid flows, structural deflections or vibrations in buildings, and the surface motion of a beating heart in minimally invasive surgery [3].

The *main challenge* in the estimation of *distributed systems* by means of a sensor network is that the individual nodes are able to measure the distributed physical quantity only at discrete time steps and discrete spatial coordinates. That means, no information between nodes and measurement steps is available.

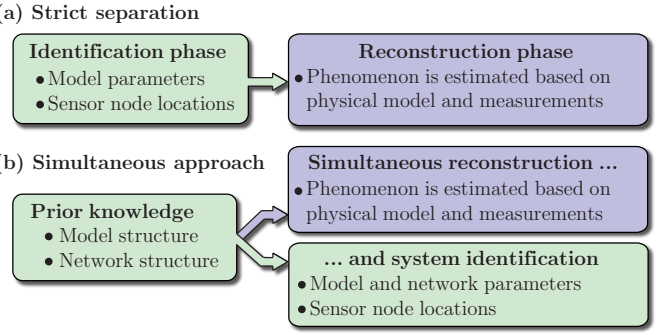

Figure 1. Procedural methods for the state and parameter estimation of distributed systems by means of a sensor network. **(a)** Strict separation of the parameter estimation (*identification phase*) and the state estimation (*reconstruction phase*). **(b)** *Simultaneous reconstruction* of the entire distributed system *and identification* of unknown parameters.

In the literature, various techniques can be found for the interpolation and extrapolation of a random field, i.e., the reconstruction of the entire distributed system. The most popular approach, especially in the geostatistic field, is the *Kriging interpolation* [4]. This approach is based on a stochastic model of the spatial dependency in terms of either variograms or mean and variance functions. Since the technique relies solely on the measured data available, the interpolation only of a realization of the distributed system is possible. That means, uncertainties in the measurements cannot be sufficiently considered. Another serious disadvantage is that additional background knowledge, such as knowledge of the physical characteristics of the distributed phenomena cannot be considered. In the statistical community the same approach is also known as *Gaussian process regression* [5], [6].

However, by exploiting *additional background information* about the *physical characteristics of the phenomenon* in form of a mathematical model, more accurate estimates can be derived; especially between the individual nodes [7].

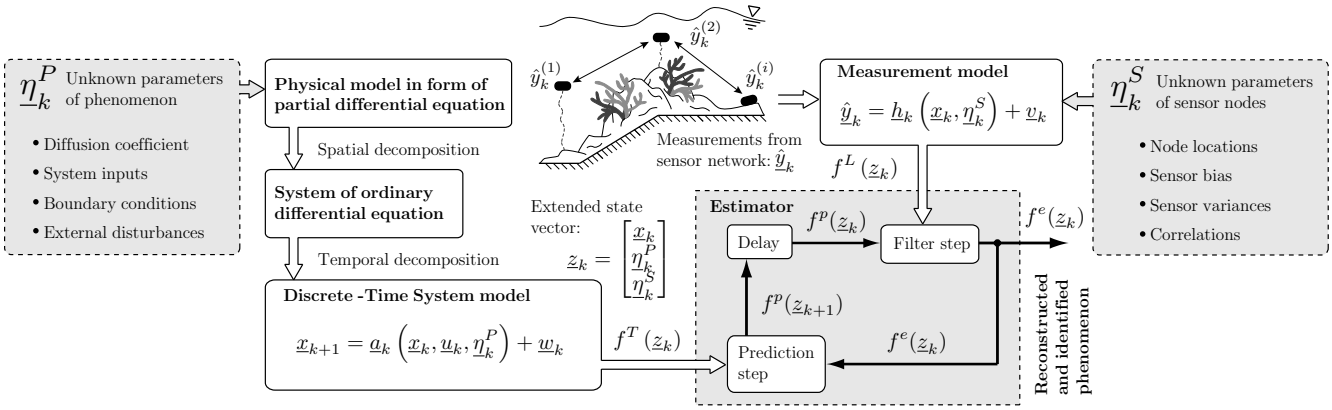

Figure 2. Overview and components of the procedure for model-based simultaneous state and parameter estimation of distributed systems.

Furthermore, by means of the model-based approach, *non-measureable* quantities can be identified, and thus additional information about the phenomenon can be obtained, such as material properties, sources of chemical concentrations or leakages. That means, one of the most important issues concerning distributed systems is the *parameter estimation*, also refered to as *parameter identification* or *inverse problem*. The main goal is the estimation of parameters in the system model from observed measurements such that the predicted state is close to the observations. The identification of such system parameters becomes even more essential for sensor network applications in harsh and unknown environments, and for unpredictable variations of the distributed system to be monitored. In Fig. 1, procedural methods for the state and parameter estimation of distributed systems are visualized: **(a)** the strict separation of the parameter estimation from the state estimation and **(b)** the simultaneous estimation approach.

For the estimation of distributed systems, the conversion of the mathematical description into a finite-dimensional model was proposed by various authors [8]–[10]. The main challenge is that the simultaneous state and parameter estimation of distributed systems leads to a *high-dimensional strongly nonlinear* estimation problem. To cope with this difficulty, special estimators based on linearizations at consecutive state trajectories [11] or linearization of the system description [12] were employed. Due to the estimation based on a linearized model, accurate results and convergence are not assured.

Fortunately, the finite-dimensional model for the simultaneous state and parameter estimation of distributed systems includes a large substructure with linear equations subject to Gaussian noise. In this case, a *decomposition* of the entire estimation problem into a (conditionally) *linear* and a *nonlinear* problem allows for an overall more efficient estimation process and more accurate results. The novelty of this paper is the exploitation of this linear substructure for the estimation of distributed system by means of a sensor network.

There are several methods to solve the combined linear/nonlinear estimation problem. The marginalized particle filter [13], [14] integrates over the linear subspace in order to reduce the dimensionality of the state space. Based on this marginalization, the standard particle filter [15] is extended by applying the Kalman filter to find the optimal estimate for the linear subspace (which is associated with the respective individual particles). Although the marginalized filter certainly improves the performance in comparison with the standard particle filter, some drawbacks still remain. For instance, in order to avoid effects like sample degeneration and impoverishment, special measures have to be taken. More important, it does not provide a measure on how well the true joint density is represented by the estimated one.

In this paper, we present a novel estimator for the simultaneous state and parameter estimation of distributed sytems. There are two key features leading to a significantly improved estimation result: **(a)** the application of a special kind of density for decomposing the estimation problem, and **(b)** a systematic approximation method leading to (close to) optimal estimation results. To be more specific, as a density representation a so-called *sliced Gaussian mixture density* is employed. The simultaneous state and parameter estimation based on such a density representation makes a systematic estimation approach feasible for large-area distributed systems. Furthermore, the uncertainties occuring in the system and arising from noisy measurements are considered by an integrated treatment. By means of the model-based approach, it is possible to identify and track unpredictable variations both of the distributed system and of the sensor network itself.

The remainder of this paper is structured as follows: Section II contains a rigorous formulation of the problem and challenges of the simultaneous state and parameter estimation of *distributed systems*. Section III is devoted to the spatial and temporal discretization allowing the conversion of the distributed system into a system description in state space form. It turns out that the parameter identification of such systems usually leads to a high-dimensional nonlinear estimation problem, however, with a linear substructure. Accordingly, in Section IV, we introduce a novel estimator, the so-called *Sliced Gaussian Mixture Filter* (SGMF), exploiting (conditionally) linear substructures in general nonlinear systems. In Section V, the performance of the proposed simultaneous estimation approach is demonstrated by means of simulation results.

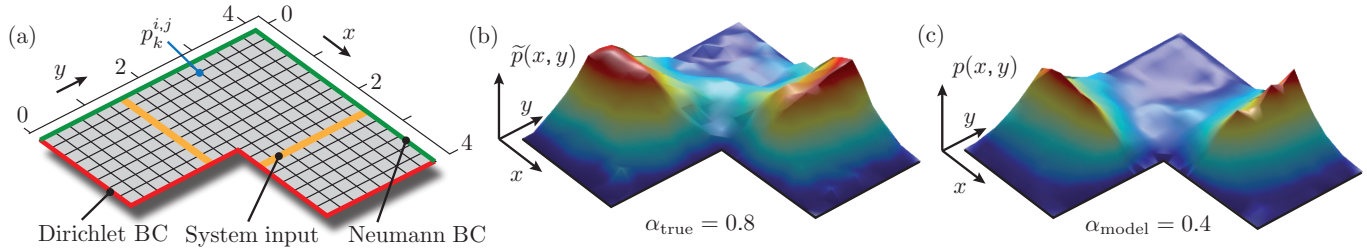

Figure 3. Exemple of a two-dimensional distributed system: **(a)** Visualization of the considered two-dimensional L-shaped domain, assumed boundary conditions, and system inputs. **(b)** Realization $\widetilde{\boldsymbol{p}}(x,y)$ of the distributed system with assumed true diffusion coefficient $\alpha_{\text{true}} = 0.8$ at time step $k = 100$. **(c)** Estimation result $\boldsymbol{p}(x,y)$ derived with standard Kalman filter based on nominal diffusion coefficient $\alpha_{\text{model}} = 0.4$.

II. Problem Formulation

The main goal is to design an efficient algorithm for the *simultaneous state and parameter* estimation of distributed systems. The evolution of a large number of physical, physiological, and even ecological systems can be described in terms of a set of partial differential equations.

In this paper, we consider only *two-dimensional linear partial differential equations* for simplicity and brevity, although similar expressions can be found for the multi-dimensional case. In its most general form, the two-dimensional linear partial differential equation without cross-derivatives is given by

$$\mathbb{L}\left(\boldsymbol{p}(\underline{z},t), \boldsymbol{s}(\underline{z},t), \frac{\partial \boldsymbol{p}}{\partial t}, \ldots, \frac{\partial^i \boldsymbol{p}}{\partial t^i}, \nabla \boldsymbol{p}, \ldots, \nabla^j \boldsymbol{p}\right) = 0 \quad , \quad (1)$$

where $\boldsymbol{p}\,(\underline{z},t)$ denotes the *distributed state* of the physical system at time t and location $\underline{z} = [x,y]$ and the operator ∇^j is defined as $\nabla^j := \frac{\partial^j}{\partial x^j} + \frac{\partial^j}{\partial y^j}$. Furthermore, the source term $\boldsymbol{s}(\underline{z},t)$, the distributed state $\boldsymbol{p}(\underline{z},t)$, and its derivatives are related by a linear operator denoted by $\mathbb{L}(\,\cdot\,)$.

The aforementioned partial differential equation (1) can be regarded as the *infinite*-dimensional form of the distributed system. However, the application of a Bayesian approach for the state and parameter estimation based on such system description is a challenging task. For that reason, the partial differential equation (1) is usually converted into a *finite*-dimensional system in state space form.

Due to the nonlinear relationship between the distributed system state $\boldsymbol{p}(\underline{z},t)$ and the unknown parameter vector $\underline{\boldsymbol{\eta}}^P(\underline{z},t)$ of the distributed system (1), the conversion leads to a *nonlinear* finite-dimensional system model according to

$$\underline{\boldsymbol{x}}_{k+1} = \underline{a}_k\left(\underline{\boldsymbol{x}}_k, \underline{\boldsymbol{\eta}}_k^P, \underline{\hat{u}}_k\right) + \underline{\boldsymbol{w}}_k^x \quad , \quad (2)$$

where $\underline{\boldsymbol{x}}_k$ contains the converted states characterizing the state of the distributed system, $\underline{\hat{u}}_k$ denotes the system input, and $\underline{\boldsymbol{w}}_k^x$ represents the system uncertainties. The parameter vector $\underline{\boldsymbol{\eta}}_k^P$ contains all the *unknown parameters* to be identified in the system model, such as unpredictable variations of physical constants. In addition, unknown constraints at the boundary of the distributed system and unknown system inputs could be considered in the parameter vector $\underline{\boldsymbol{\eta}}_k^P$. The system model and examples of parameters to be estimated by a sensor network are visualized in Fig. 2.

Besides, there is a measurement model describing the physical properties of the sensor network itself. In general, the measurements $\underline{\hat{y}}_k$ are related *nonlinearly* to the state vector $\underline{\boldsymbol{x}}_k$, according to

$$\underline{\hat{y}}_k = \underline{h}_k\left(\underline{\boldsymbol{x}}_k, \underline{\boldsymbol{\eta}}_k^S\right) + \underline{\boldsymbol{v}}_k \quad , \quad (3)$$

where $\underline{\boldsymbol{v}}_k$ is the uncertainty in the measurement model. The parameter vector $\underline{\boldsymbol{\eta}}_k^S$ contains *unknown parameters* to be identified in the measurement model. Sensor bias and sensor variances, for example, could be included in the unknown parameter vector $\underline{\boldsymbol{\eta}}_k^S$ for the purpose of tracking physical wear of the sensor nodes. Furthermore, one could imagine to collect unknown node locations and correlations in the parameter vector $\underline{\boldsymbol{\eta}}_k^S$, see Fig. 2.

It is shown that for the simultaneous state and parameter estimation of distributed systems, the nonlinear system function $\underline{a}_k\,(\,\cdot\,)$ includes a high–dimensional linear substructure. This allows a decomposition of the total state vector $\underline{\boldsymbol{z}}_k$ to be estimated into two substate vectors,

$$\underline{\boldsymbol{z}}_k = \begin{bmatrix}(\underline{\boldsymbol{x}}_k)^T & (\underline{\boldsymbol{\eta}}_k)^T\end{bmatrix}^T \quad , \quad (4)$$

with the high–dimensional state vector $\underline{\boldsymbol{x}}_k \in \mathbb{R}^r$ (characterizing the conditionally linear system) and the lower–dimensional parameter vector $\underline{\boldsymbol{\eta}}_k \in \mathbb{R}^s$ (characterizing the nonlinear part of the system).

For the estimation of the total state vector $\underline{\boldsymbol{z}}_k$, the decomposition into a state vector $\underline{\boldsymbol{x}}_k$ and parameter vector $\underline{\boldsymbol{\eta}}_k$ is exploited for the derivation of a more efficient estimator than nonlinear estimators operating on the entire vector $\underline{\boldsymbol{z}}_k$. This decomposition of the estimation problem into a *linear* and a *nonlinear* problem is mainly achieved by a novel density representation, the so-called *sliced Gaussian mixture density*, and the systematic approximation of arbitrary densities by this representation.

Here, we emphasize that the estimation approach introduced in this paper is not restricted to the application to distributed systems. This approach can always be applied when the simultaneous state and parameter estimation task of a general dynamic system leads to a conditionally linear system description. In the case of distributed systems, the linear substructure can be significantly larger compared to the nonlinear substructure.

III. Simultaneous State and Parameter Estimation of Distributed Systems

In this section, we explain a method for the state and parameter estimation by means of discrete space-time measurements performed by a sensor network. The methods introduced here can be applied to the general case of linear partial differential equations (1) and could even be extended to the multi-dimensional case in a straightforward fashion. However, we restrict our attention to a certain distributed system, the so-called diffusion equation.

Example 1 (Diffusion equation)
Throughout this paper, we consider the following two-dimensional linear partial differential equation,

$$\mathbb{L}\left(\boldsymbol{p}(\underline{z},t)\right) = \frac{\partial \boldsymbol{p}(\underline{z},t)}{\partial t} - \alpha(\underline{z},t)\,\nabla^2 \boldsymbol{p}(\underline{z},t) - \boldsymbol{s}(\underline{z},t) = 0 \ , \quad (5)$$

where the diffusion coefficient $\alpha(\underline{z},t)$ could be both time and space varying. The aim is the estimation of the solution $\boldsymbol{p}(\underline{z},t)$ and the identification of the unknown parameter $\boldsymbol{\eta}^P = \alpha(\underline{z},t)$ in a *simultaneous fashion*.

A. Conversion of the Distributed System

The model-based state estimation of distributed systems based on a distributed-parameter description is quite complex. The reason is that for a Bayesian estimation method usually a lumped-parameter system description is used. To cope with this problem, the system description is converted from a distributed-parameter into a lumped-parameter form. This conversion can be achieved by methods for solving partial differential equations, such as finite-difference method [16], the finite-element method, modal analysis [3] and finite-spectral method [17].

The simplest method for the spatial and temporal discretization of distributed system is the finite-difference method. In order to solve the partial differential equation (5), the derivatives need to be approximated with finite differences according to

$$\frac{\partial \boldsymbol{p}(\underline{z},t)}{\partial t} \approx \frac{\boldsymbol{p}_{k+1}^{i,j} - \boldsymbol{p}_{k}^{i,j}}{\Delta t} \ ,$$

$$\nabla^2 \boldsymbol{p}(\underline{z},t) \approx \frac{\boldsymbol{p}_{k}^{i+1,j} + \boldsymbol{p}_{k}^{i-1,j} + \boldsymbol{p}_{k}^{i,j+1} + \boldsymbol{p}_{k}^{i,j-1} - 4\boldsymbol{p}_{k}^{i,j}}{\Delta h^2} \ ,$$

where Δt is the sampling time and Δh denotes the spatial sampling period. The superscript i,j and subscript k in $\boldsymbol{p}_k^{i,j}$ denote the value of the distributed system at discretization node (i,j) and at time step k, respectively.

Example 2 (Rectangular solution domain)
In this example, we illustrate the structure of the converted diffusion equation (5) derived by means of the finite-difference method. Here, we assume to have a rectangular solution domain with respective boundary conditions. Then, the conversion of the distributed system (5) results in the following system matrix $\mathbf{A}_k \in \mathbb{R}^{m^2 \times m^2}$,

$$\mathbf{A}_k = \frac{\alpha_k\,\Delta t}{\Delta h^2} \begin{bmatrix} \widetilde{\mathbf{A}}_m & \mathbf{I}_m & \mathbf{0} & \dots & \mathbf{0} \\ \mathbf{I}_m & \widetilde{\mathbf{A}}_m & \mathbf{I}_m & \dots & \mathbf{0} \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ \mathbf{0} & \dots & \mathbf{I}_m & \widetilde{\mathbf{A}}_m & \mathbf{I}_m \\ \mathbf{0} & \dots & \mathbf{0} & \mathbf{I}_m & \widetilde{\mathbf{A}}_m \end{bmatrix} + \mathbf{I}_{m^2} \ ,$$

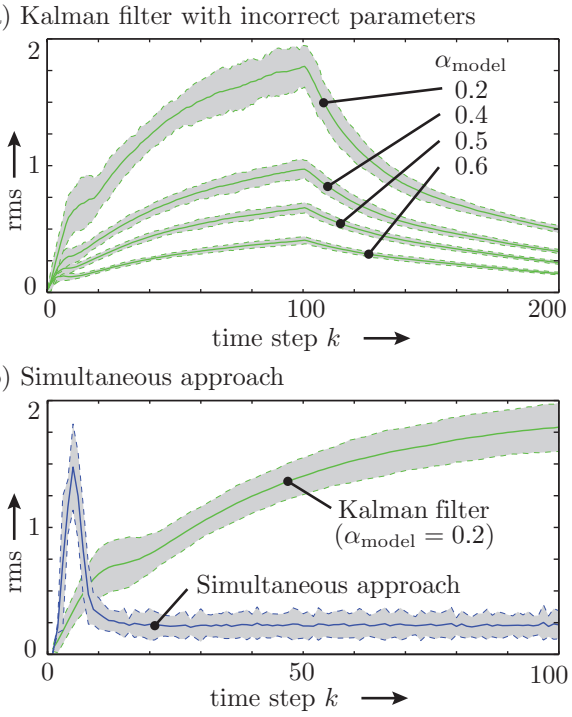

Figure 4. Root mean square error $\hat{e}_k$ and error variance C_k^{rms} for the estimated distributed system for 50 Monte Carlo simulation runs. The true parameter α_{true} is given by $\alpha_{\text{true}} = 0.8$. **(a)** Kalman filter (green) based on various incorrect parameters $\alpha_{\text{model}} = \{0.2,\ 0.4,\ 0.5,\ 0.6\}$. **(b)** Kalman filter (green) based on incorrect model parameter ($\alpha_{\text{model}} = 0.2$) and simultaneous state and parameter estimation approach (blue). Thanks to simultaneous approach the performance of the estimation result can be significantly improved.

where $\mathbf{I}_{m^2} \in \mathbb{R}^{m^2 \times m^2}$ represents the identity matrix and the sub-matrices $\widetilde{\mathbf{A}}_m \in \mathbb{R}^{m \times m}$ are given by

$$\widetilde{\mathbf{A}}_m = \begin{bmatrix} -4 & 1 & 0 & \dots & 0 \\ 1 & -4 & 1 & \dots & 0 \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ 0 & \dots & 1 & -4 & 1 \\ 0 & \dots & 0 & 1 & -4 \end{bmatrix} \ .$$

The state of the distributed system is characterized by the state vector $\underline{\boldsymbol{x}}_k = [\boldsymbol{p}_k^{1,1}, \dots, \boldsymbol{p}_k^{1,m}, \dots, \boldsymbol{p}_k^{m,m}]$. For the conversion of the entire distributed system, the input function $\boldsymbol{s}(\underline{z},t)$ needs to be discretized in the same way as the system state. This leads to the input vector $\underline{\boldsymbol{u}}_k = [\boldsymbol{s}_k^{1,1}, \dots, \boldsymbol{s}_k^{1,m}, \dots, \boldsymbol{s}_k^{m,m}]$. The input matrix $\mathbf{B}_k$ relating the input $\underline{\boldsymbol{u}}_k$ of the distributed system to its state vector $\underline{\boldsymbol{x}}_k$ is given by a diagonal matrix with the sampling time Δt as the diagonal entries, according to $\mathbf{B}_k = \text{diag}\,\{\Delta t, \dots, \Delta t\}$. The detailed description of the boundary conditions, such as Dirichlet boundary condition and Neuman boundary condition, and how these need to be considered during the conversion process is omitted in this paper; instead we refer to [16].

B. System and Measurement Equation

In the previous section, we presented the spatial and temporal discretization allowing the conversion of the distributed-parameter system into a lumped-parameter form. The application of the aforementioned methods to *linear* partial differential equations (1) always results in a linear system of equations for the state vector characterizing the distributed

system [8], [12]. Adding noise terms and modelling error terms leads to the following system equation

$$\underline{\boldsymbol{x}}_{k+1} = \mathbf{A}_k \underline{\boldsymbol{x}}_k + \mathbf{B}_k \left(\underline{\hat{u}}_k + \underline{\boldsymbol{w}}_k^x \right) \quad , \tag{6}$$

where the structure of the system matrix $\mathbf{A}_k$ and the input matrix $\mathbf{B}_k$ merely depends on the applied conversion method.

The measurement equation providing a mapping of the finite-dimensional state vector $\underline{\boldsymbol{x}}_k$ to the individual discrete-time measurements $\underline{\hat{y}}_k$ can always be stated in a linear form according to

$$\underline{\hat{y}}_k = \mathbf{H}_k \underline{\boldsymbol{x}}_k + \underline{\boldsymbol{v}}_k \quad , \tag{7}$$

independent on the used method for the conversion. The measurement matrix $\mathbf{H}_k$ is defined on the basis of geometric relations between the state vector $\underline{\boldsymbol{x}}_k$ and the sensor locations. Thus, it essentially depends on the shape functions used for the spatial discretization. Here, we refer to our previous research work [7] for a more detailed description on the structure and derivation of the measurement matrix for distributed systems.

C. Reconstruction with Incorrect Model Parameters

In general, depending on the structure of the system model and the measurement model, i.e., being linear or nonlinear, an appropriate estimator has to be chosen in order to estimate the state characterizing the distributed system. Due to the fact that both system equation (6) and the measurement equation (7) are linear, it is sufficient to use the linear Kalman filter to obtain the best possible estimate.

It is well known that the Kalman filter requires a rather precise model of the system under consideration and a precisely known noise statistics. If any of these assumptions is violated, the performance of the filter estimates quickly degrades. However, in many cases the real system deviates from the nominal model. The resulting degradation leading to poor performance is illustrated in the next example.

Example 3 (System model with incorrect parameters)
In this example, we consider the two-dimensional diffusion equation (5) in an L-shaped solution domain with respective boundary conditions and system inputs as visualized in Fig. 3 (a). The distributed system is converted into a lumped-parameter system described by $m = 243$ state variables $\boldsymbol{p}_k^{i,j}$. The nominal parameter values for the system model (6) are given by

$$\Delta t = 0.01 \quad , \quad \Delta h = 0.5 \quad , \quad \alpha_{\text{true}} = 0.8 \quad ,$$

where α_{true} is assumed to be the true parameter. The system input $\hat{u}_k^i$ at the locations visualized in Fig. 3 (a) are given by

$$\hat{u}_k^i = \begin{cases} 8 & \text{for} \quad 0 \leq k \leq 100 \\ 0 & \text{for} \quad 100 \leq k \leq 200 \end{cases} .$$

Furthermore, there are *sensor nodes* at every *discretization node* (i, j) with the measurement noise variance $C_k^v = 0.02$. At every time step, 20 randomly chosen sensor nodes are performing a measurement step in order to estimate the entire state of the distributed system. The state estimation of the distributed system was performed on the basis of a Kalman filter with the nominal parameter set for the diffusion coefficient α according to

$$\alpha_{\text{model}} = \{0.2,\ 0.4,\ 0.5,\ 0.6\} \quad ,$$

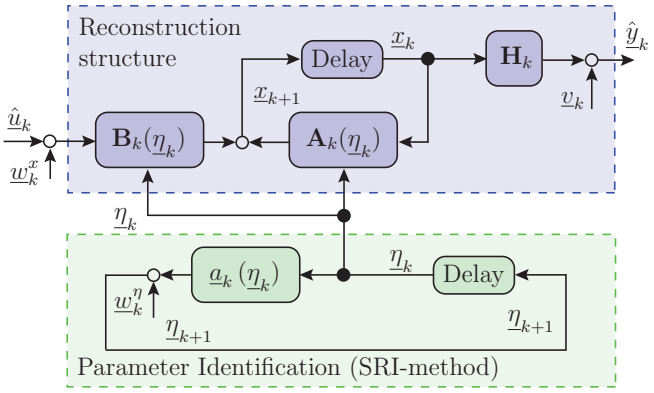

Figure 5. Visualization of a dynamic system with a linear substructure. The parameter $\underline{\boldsymbol{\eta}}_k$ characterizes the system matrix $\mathbf{A}_k$ and the input matrix $\mathbf{B}_k$, and thus the dynamic behavior of the conditionally linear system.

with the true parameter $\alpha_{\text{true}} = 0.8$. For each parameter value, 50 independent Monte Carlo simulation runs have been performed, resulting in $n = 50$ true realizations $\widetilde{\underline{x}}_k^i$ of the state vector. The simulation results are shown in Fig. 3 and Fig. 4.

In Fig. 3 **(b)**, **(c)** a true realization of the distributed system and the respective estimation result based on a deviated parameter of $\alpha_{\text{model}} = 0.4$ is depicted. It is obvious that the estimated distributed state $\boldsymbol{p}(x, y)$ based on incorrect model parameters strongly deviates from the true realization $\widetilde{\boldsymbol{p}}(x, y)$.

The root mean square error and the error variance is approximated by calculating the average according to

$$\hat{e}_k^2 \approx \frac{1}{n \cdot m} \sum_{i=1}^{n} \| \widetilde{\underline{x}}_k^i - \hat{\underline{x}}_k^i \| \, , \quad C_k^{\text{rms}} \approx \frac{1}{n-1} \sum_{i=1}^{n} \left(e_k^i - \hat{e}_k \right)^2 \quad ,$$

where $\hat{\underline{x}}_k^i$ denotes the mean of the estimated state vector. The root mean square error $\hat{e}_k$ and error variance C_k^{rms} for each nominal parameter value are shown in Fig. 4 **(a)**. The more the nominal parameters deviate from the true parameters, the more the performance of the estimation results degrades.

In Fig. 4 **(b)**, the error between the Kalman filter based on incorrect parameters and the simultaneous state and parameter estimator is compared (see Sec. III-D and Sec. IV). It is obvious that thanks to the simultaneous state and parameter estimation the performance can be significantly increased.

D. State and Parameter Estimation of Distributed Systems

For the *simultaneous state and parameter estimation* of distributed systems, the unknown parameter vector $\underline{\boldsymbol{\eta}}_k$ is treated as an additional part of the state vector. By this means, conventional estimation techniques can be used to estimate the parameter and states simultaneously (also called *joint estimation*).

Hence, an *augmented state vector* $\underline{\boldsymbol{z}}_k$ containing the system state $\underline{\boldsymbol{x}}_k$ and the additional unknown parameters is defined by

$$\underline{\boldsymbol{z}}_k := \begin{bmatrix} \underline{\boldsymbol{x}}_k \\ \underline{\boldsymbol{\eta}}_k \end{bmatrix} \quad .$$

In the case of system identification, the augmentation results in the following *augmented system model*

$$\begin{bmatrix} \underline{\boldsymbol{x}}_{k+1} \\ \underline{\boldsymbol{\eta}}_{k+1} \end{bmatrix} = \begin{bmatrix} \mathbf{A}_k(\underline{\boldsymbol{\eta}}_k)\underline{\boldsymbol{x}}_k + \mathbf{B}_k(\underline{\boldsymbol{\eta}}_k)\hat{\underline{u}}_k \\ \underline{a}_k(\underline{\boldsymbol{\eta}}_k) \end{bmatrix} + \begin{bmatrix} \mathbf{B}_k(\underline{\boldsymbol{\eta}}_k)\underline{\boldsymbol{w}}_k^x \\ \underline{\boldsymbol{w}}_k^\eta \end{bmatrix} \quad , \quad (8)$$

and measurement model

$$\hat{\underline{y}}_k = \mathbf{H}_k\, \underline{\boldsymbol{x}}_k + \underline{\boldsymbol{v}}_k \quad , \quad (9)$$

where the nonlinear function $\underline{a}_k(\cdot)$ describes the dynamic behavior of the parameter $\underline{\boldsymbol{\eta}}_k$ to be estimated. The structure of the augmented system is visualized in Fig. 5. It can be clearly seen that the parameter $\underline{\boldsymbol{\eta}}$ characterizes the system matrix $\mathbf{A}_k$ and the input matrix $\mathbf{B}_k$, and thus the dynamic behavior of the conditionally linear system.

In the case of the simultaneous state and parameter estimation of distributed systems, the augmented system model is *nonlinear* in the augmented state $\underline{\boldsymbol{z}}_k$. This is mainly due to the multiplication of $\mathbf{A}_k(\underline{\boldsymbol{\eta}}_k)$ containing the unknown parameter $\underline{\boldsymbol{\eta}}_k$ and the system state $\underline{\boldsymbol{x}}_k$. However, the system model (8) contains a high-dimensional linear substructure, which can be exploited by the application of a more efficient estimator. In the following section, we describe a novel estimator – *Sliced Gaussian Mixture Filter* (SGMF) – allowing the decomposition of the estimation problem.

IV. Sliced Gaussian Mixture Filter (SGMF)

For the exploitation of linear substructures in general nonlinear systems, we introduced in our previous research work [18] a systematic estimator, the so-called *Sliced Gaussian Mixture Filter* (SGMF). There are two key features leading to a significantly improved estimation result compared to other state of the art estimation approaches.

- **Novel density representation:** The utilization of a special kind of density allows the decomposition of the general estimation problem into a linear and nonlinear problem. To be more specific, as a density representation the so-called *sliced Gaussian mixture density* is employed for the simultaneous state and parameter estimation of distributed systems.
- **Systematic approximation:** The systematic approximation of the density resulting from the estimation update leads to (close to) optimal approximation results. Thus, less parameters for the density representation are necessary and a measure for the approximation performance is provided.

Despite the high-dimensional nonlinear character, the systematic approach of the simultaneous state and parameter estimation for large-area distributed phenomena is feasible thanks to the decomposition based on sliced Gaussian mixture density. Furthermore, the uncertainties occuring in the mathematical system description and arising from noisy measurements are considered by an integrated treatment.

The *Sliced Gaussian Mixture Filter* basically consists of three steps: the *decomposition* of the estimation problem, the utilization of an *efficient update*, and the *reapproximation* of the density representation.

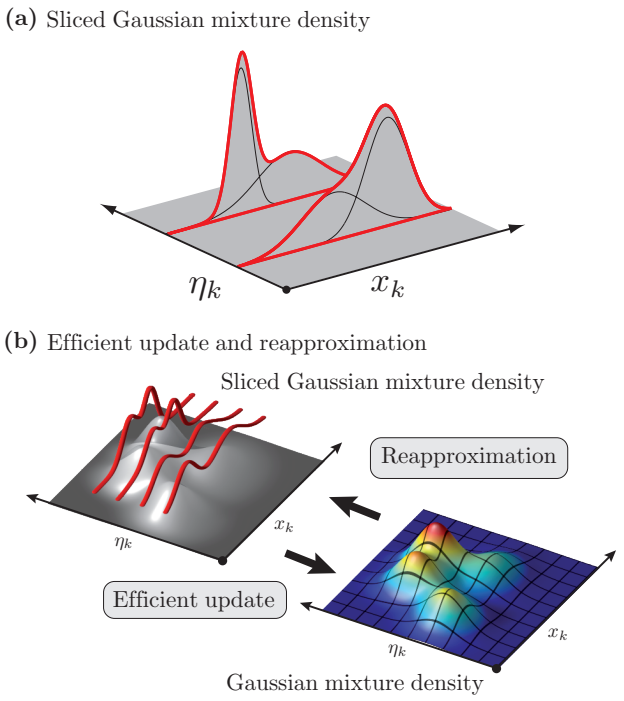

Figure 6. **(a)** The estimator is based on *sliced Gaussian mixture densities* consisting of a Gaussian mixture in $\underline{\boldsymbol{x}}_k$ subspace and Dirac mixture in $\underline{\boldsymbol{\eta}}_k$ subspace. By this means, the estimation problem can be decomposed into a linear and a nonlinear problem. **(b)** Procedure of the simultaneous estimation of the state $\underline{\boldsymbol{x}}_k$ and the parameter $\underline{\boldsymbol{\eta}}_k$. The estimation result of the Sliced Gaussian Mixture Filter leads to posterior Gaussian mixture density, which is then reapproximated.

*a) **Decomposition**:* The nonlinear high-dimensional estimation problem is decomposed into a *linear* high-dimensional problem (state estimation) and a *nonlinear* low-dimensional problem (parameter estimation). This can be achieved by means of the *sliced Gaussian mixture density*, see Fig. 6 **(a)**. The sliced Gaussian mixture density $f(\underline{x}_k, \underline{\eta}_k)$ is represented by a Dirac mixture in the *nonlinear* subspace $\underline{\boldsymbol{\eta}}_k$ (parameter space) and Gaussian mixture in the linear subspace $\underline{\boldsymbol{x}}_k$ (state space),

$$f(\underline{x}_k, \underline{\eta}_k) = \sum_{i=1}^{M} \alpha_k^i \delta\left(\underline{\eta}_k - \underline{\xi}_k^i\right) \sum_{j=1}^{N^i} \beta_k^{ij} \mathcal{N}\left(\underline{x}_k - \underline{\mu}_k^{ij}, \mathbf{C}_k^{ij}\right). \quad (10)$$

where $\underline{\xi}_k^i \in \mathbb{R}^s$ can be regarded as the position of the slices of the individual sliced Gaussian mixture density $f(\underline{x}_k, \underline{\eta}_k)$. The β_k^{ij}, $\underline{\mu}_k^{ij} \in \mathbb{R}^r$, and $\mathbf{C}_k^{ij} \in \mathbb{R}^{r\times r}$ denote the weights, means, and covariance matrices of the j-th component of the Gaussian mixture density of the i-th slice.

*b) **Efficient update**:* Thanks to the novel density representation (10) and the structure of the augmented system model (8) an overall more efficient estimation update can be derived. The proof can be found in our previous research work [18]. By this means, the predicted density $\widetilde{f}^p$ results in a Gaussian mixture both in linear subspace $\underline{\boldsymbol{x}}_k$ and nonlinear

Conditionally linear subspace
$\gamma_k^{ij} := \mathcal{N}\left(\hat{\underline{y}}_k - \mathbf{H}_k\underline{\mu}_k^{pij}, \mathbf{H}_k\mathbf{C}_k^{pij}\mathbf{H}_k{}^T + \mathbf{C}_v\right)$
$\underline{\mu}_k^{eij} := \underline{\mu}_k^{pij} + \mathbf{K}\left(\hat{\underline{y}}_k - \mathbf{H}_k\underline{\mu}_k^{pij}\right)$
$\mathbf{C}_k^{eij} := \mathbf{C}_k^{pij} - \mathbf{K}\mathbf{H}_k\mathbf{C}_k^{pij}$
with $\mathbf{K} := \mathbf{C}_k^{pij}\mathbf{H}_k{}^T\left(\mathbf{C}_v + \mathbf{H}_k\mathbf{C}_k^{pij}\mathbf{H}_k{}^T\right)^{-1}$

Table I
FILTER STEP: PARAMETERS OF ESTIMATED DENSITY.

Nonlinear subspace	Conditionally linear subspace
$\underline{\xi}_{k+1}^{pi} := \underline{a}_k\left(\underline{\xi}_k^{ei}\right)$	$\underline{\mu}_{k+1}^{pij} := \mathbf{A}_k^i\underline{\mu}_k^{eij} + \mathbf{B}_k^i\underline{\hat{u}}_k$
$\mathbf{C}_w^n$	$\mathbf{C}_{k+1}^{pij} := \mathbf{A}_k^i\mathbf{C}_k^{eij}\mathbf{A}_k^i{}^T + \mathbf{C}_w^l$

Table II
PREDICTION STEP: PARAMETERS OF PREDICTED DENSITY.

subspace $\underline{\boldsymbol{\eta}}_k$

$$\widetilde{f}^p(\underline{x}_{k+1}, \underline{\eta}_{k+1}) = c \cdot \sum_{i=1}^{M}\sum_{j=1}^{N^i} \alpha_k^i\beta_k^{ij}\gamma_k^{ij} \cdot \mathcal{N}\left(\underline{\eta}_{k+1} - \underline{\xi}_{k+1}^{pi}, \mathbf{C}_w^n\right)\mathcal{N}\left(\underline{x}_{k+1} - \underline{\mu}_{k+1}^{pij}, \mathbf{C}_{k+1}^{pij}\right), \quad (11)$$

where the mean and covariance matrices in linear subspace $\underline{\boldsymbol{x}}_k$ are calculated by applying the standard Kalman prediction and measurement step. The mean in nonlinear subspace $\underline{\boldsymbol{\eta}}_k$ is derived by simply repositioning the density slices according to the nonlinear system equation (8), see Table I and II.

*c) **Reapproximation**:* The estimation based on the sliced Gaussian mixture density leads to a density representation consisting of *Gaussian mixtures* in all subspaces. In order to bound the complexity, the resulting density needs to be reapproximated by means of the *sliced Gaussian mixture density*. There are several approaches to perform this approximation. One possible approach for the approximation is to derive the location of the density slices by only considering the marginal density $\widetilde{f}^p(\underline{\eta}_{k+1}^S)$. The approximation of arbitrary marginal densities by Dirac mixture densities can be achieved by: *batch approximation* [19] or *sequential approximation* [20].

The *batch approximation* is an efficient solution procedure for arbitrary density functions on the basis of homotopy continuation (Progressive Bayes). This procedure results in an optimal solution. The *sequential approximation* is based on inserting one component of the density slices at a time. In the scalar case every slice corresponds to an interval in the nonlinear subspace and approximates the true marginal density only in the corresponding interval. Then, based on the splitting of the intervals and their respective slices arbitrary densities can be approximated [18].

After the approximation of the marginal density $\widetilde{f}^p(\underline{\eta}_{k+1}^S)$ in the nonlinear subspace, the Dirac approximation is extended to a sliced Gaussian mixture representation over the entire sample space. Basically, this is achieved by evaluating the Gaussian mixture density $\widetilde{f}^p(\underline{x}_{k+1}, \underline{\eta}_{k+1}^S)$ at every Dirac position, i.e., at every slice position. This leads to a sliced Gaussian mixture density (10), which can be used for the next processing step. A more detailed description can be found in [18].

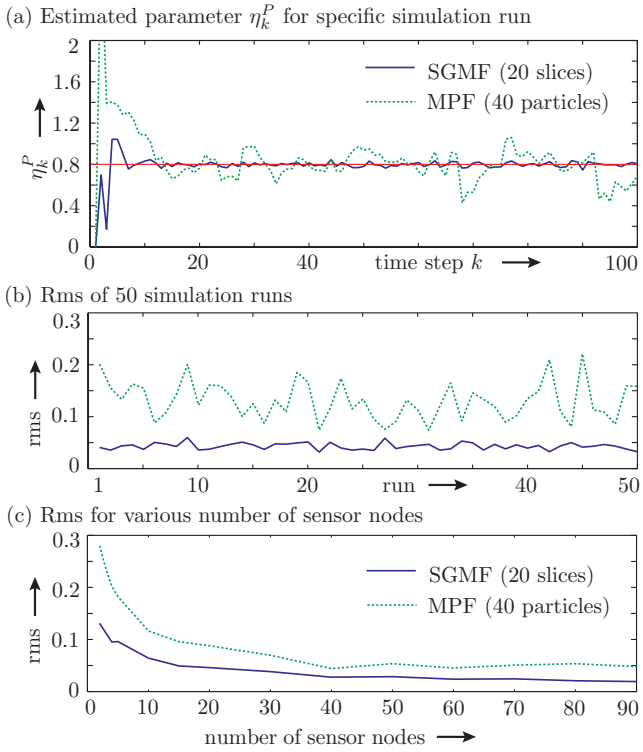

Figure 7. Comparison of the Sliced Gaussian Mixture Filter (SGMF) using 20 slices and the marginalized particle filter (MPF) using 40 particles for assumed true parameter $\eta_k^P = 0.8$. **(a)** Mean estimation $\hat{\eta}_k^P$ of the parameter for an example simulation run. **(b)** Root mean square error (rms) of 50 simulation runs. **(c)** Rms averaged of all simulation runs for various number of sensor nodes.

V. SIMULATION RESULTS

In this section, the performance of the simultaneous state and parameter estimation of distributed systems based on the Sliced Gaussian Mixture Filter is demonstrated by means of simulation results. In particular, the accuracy of the identified parameter vector $\boldsymbol{\eta}_k$ characterizing the distributed system is investigated in comparison to another nonlinear estimation method. The following distributed system is considered:

Example 4 (Considered distributed system)
In this simulation, we consider the two-dimensional diffusion equation on a L-shaped solution domain, see Fig. 3 **(a)**, and with assumed boundary conditions and system inputs as described in Example **1–3**. The aim is the simultaneous estimation of the distributed system and the unknown diffusion coefficient $\boldsymbol{\eta}_k^P = \boldsymbol{\alpha}_k$, where the true parameter is given by $\boldsymbol{\alpha}_{\text{true}} = 0.8$. The system noise term for the individual discretization nodes is assumed to be $C_k^w = 0.01$. Furthermore, there is a sensor node at every discretization node $p_k^{i,j}$ with a measurement noise variance $C_k^v = 0.02$. At every time step, 20 randomly chosen sensor nodes are performing a measurement step. The comparison of the Sliced Gaussian Mixture Filter and the marginalized particle filter for 50 Monte Carlo runs is shown in Fig. 7.

In Fig. 7 **(a)** an example simulation run for the estimation of the parameter $\boldsymbol{\eta}_k^P$ derived by the Sliced Gaussian Mixture Filter (SGMF, 20 slices) and the marginalized particle filter [13] (MPF, 40 particles) is visualized. It is obvious that after a certain transition time the SGMF offers a nearly exact parameter estimation, while the MPF strongly jitters. The root mean square errors (rms) of all 50 runs are depicted in Fig. 7 **(b)**, where it can be clearly seen that the SGMF always outperforms the MPF. This is basically due to the systematic (non-random) positioning of the slices in the case of the SGMF, while the slices for the MPF are placed randomly. In Fig. 7 **(c)** the root mean square error averaged over all 50 simulation runs is shown for various numbers of sensor nodes. It can be seen that more sensor nodes result in a more accurate estimation of the unknown parameter $\boldsymbol{\eta}_k^P$. In this simulated case, the application of more than approximately 40 sensor nodes does not lead to further improvements of the estimated parameter $\boldsymbol{\eta}_k^P$. Furthermore, Fig. 4 **(b)** clearly shows that thanks to the simultaneous state and parameter estimation based on the Sliced Gaussian Mixture Filter the accuracy of the estimated distributed system can be increased.

VI. Conclusion and Future Works

In this paper, we introduce an efficient method for the simultaneous state and parameter estimation of *distributed systems*. The spatial and temporal decomposition of the distributed system results in a finite-dimensional model in state space form (usually characterized by a high-dimensional state vector). Hence, the augmentation of the system state with the parameter to be estimated leads to a *high-dimensional nonlinear* system description. Based on a novel density representation – *sliced Gaussian mixture density* – the linear substructure contained in the finite-dimensional model is exploited. This leads to an overall more efficient estimation process of distributed systems. The performance is demonstrated by means of simulation results and it turned out that, compared to other nonlinear estimators, the *Sliced Gaussian Mixture Filter* achieves a higher accuracy.

The application of simultaneous state and parameter estimation methods to sensor network provides novel prospects. The network is capable of estimating the *entire* state of the distributed system, identifying *non-measurable* quantities, verifying and validate the correctness of the estimation results, and adapt autonomously their algorithms and behavior to changes.

So far, the node locations were assumed to be precisely known for the estimation of distributed systems. In many real world applications, however, the node locations contain uncertainties, or even could be completely unknown. By means of the estimation method introduced in this paper, it is possible to consider this uncertainty in the node locations or localize the sensor nodes based on local observations of the distributed system. For the observation of large-area distributed systems, decentralized methods are inevitable in order to cope with high-dimensional state vectors. Hence, further decompositions both in the linear subspace and nonlinear subspace are necessary. This is left for future research work.

VII. Acknowledgments

This work was partially supported by the German Research Foundation (DFG) within the Research Training Group GRK 1194 “Self-organizing Sensor-Actuator-Networks”.

References

[1] T. Kumar, F. Zhao, and D. Shepherd, “Collaborative Signal and Information Processing in Microsensor Networks,” *IEEE Signal Processing Magazine*, vol. 19, pp. 13–14, 2002.

[2] T. C. Henderson, C. Sikorski, E. Grant, and K. Luthy, “Computational Sensor Networks,” in *Proceedings of the 2007 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS 2007)*, San Diego, USA, 2007.

[3] T. Bader, A. Wiedemann, K. Roberts, and U. D. Hanebeck, “Model-based Motion Estimation of Elastic Surfaces for Minimally Invasive Cardiac Surgery,” in *IEEE International Conference on Robotics and Automation (ICRA)*, Roma, Italia, April 2007.

[4] P. J. Curran and P. M. Atkinson, “Geostatistics and Remote Sensing,” *Progress in Physical Geography*, vol. 22, pp. 61–78, 1998.

[5] C. K. I. Williams and C. E. Rasmussen, “Gaussian Processes for Regression,” in *Proc. Conf. Advances in Neural Information Processing Systems (NIPS)*, D. S. Touretzky, M. C. Mozer, and M. E. Hasselmo, Eds., vol. 8. MIT Press, 1995.

[6] C. K. I. Williams, “Prediction with Gaussian Processes: From Linear Regression to Linear Prediction and Beyond,” *Learning in graphical models*, pp. 599–621, 1999.

[7] F. Sawo, K. Roberts, and U. D. Hanebeck, “Bayesian Estimation of Distributed Phenomena using Discretized Representations of Partial Differential Equations,” in *3rd International Conference on Informatics in Control, Automation and Robotics (ICINCO’06)*, Setubal, Portugal, Aug. 2006, pp. 16–23.

[8] D. Ucinski and J. Korbicz, “Parameter Identification of Two-Dimensional Distributed Systems,” *International Journal of Systems Science*, vol. 21, no. 12, pp. 2441–2456, 1990.

[9] J. Yin, V. L. Syrmos, and D. Y. Y. Yun, “System Identification using the Extended Kalman Filter with Applications to Medical Imaging,” in *Proceedings of the American Control Conference (ACC)*, 2000.

[10] F. Sawo, M. F. Huber, and U. D. Hanebeck, “Parameter Identification and Reconstruction Based on Hybrid Density Filter for Distributed Phenomena,” in *10th International Conference on Information Fusion (Fusion 2007)*, Quebec, Canada, Jul. 2007.

[11] H. Malebranche, “Simultaneous State and Parameter Estimation and Location of Sensors for Distributed Systems,” *International Journal of Systems Science*, vol. 19, no. 8, pp. 1387–1405, 1988.

[12] L. A. Rossi, B. Krishnamachari, and C.-C. Kuo, “Distributed Parameter Estimation for Monitoring Diffusion Phenomena Using Physical Models,” in *First Annual IEEE Communications Society Conference on Sensor and Ad Hoc Communications and Networks (SECON)*, Los Angeles, USA, 2004, pp. 460–469.

[13] T. Schön, F. Gustafsson, and P.-J. Nordlund, “Marginalized Particle Filters for Mixed Linear/Nonlinear State-Space Models,” *IEEE Transactions on Signal Processing*, vol. 53, no. 7, pp. 2279–2289, 2005.

[14] R. Chen and J. S. Liu, “Mixture Kalman Filters,” *Journal of the Royal Statistical Society*, vol. 62, no. 3, pp. 493–508, 2000.

[15] C. Andrieu and A. Doucet, “Particle Filtering for Partially Observed Gaussian State Space Models,” *Journal of the Royal Statistical Society*, vol. 64, no. 4, pp. 827 – 836, 2002.

[16] T. Chung, *Computational Fluid Dynamics*. Cambridge University Press, 2002.

[17] G. E. Karniadakis and S. Sherwin, *Spectral/hp Element Methods for Computational Fluid Dynamics*. Oxford University Press, 2005.

[18] V. Klumpp, F. Sawo, U. D. Hanebeck, and D. Fränken, “The Sliced Gaussian Mixture Filter for Efficient Nonlinear Estimation,” in *11th International Conference on Information Fusion (Fusion 2008)*, Cologne, Germany, July 2008.

[19] O. C. Schrempf and U. D. Hanebeck, “A State Estimator for Nonlinear Stochastic Systems Based on Dirac Mixture Approximations,” in *4th Intl. Conference on Informatics in Control, Automation and Robotics (ICINCO 2007)*, vol. SPSMC, Angers, France, May 2007, pp. 54–61.

[20] U. D. Hanebeck and O. C. Schrempf, “Greedy Algorithms for Dirac Mixture Approximation of Arbitrary Probability Density Functions,” in *IEEE Conference on Decision and Control (CDC 2007)*, Dec. 2007.