

Parameter Identification and Reconstruction for Distributed Phenomena Based on Hybrid Density Filter

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

Abstract—This paper addresses the problem of model-based reconstruction and parameter identification of distributed phenomena characterized by partial differential equations. The novelty of the proposed method is the systematic approach and the integrated treatment of uncertainties, which naturally occur in the physical system and arise from noisy measurements. The main challenge of accurate reconstruction is that model parameters, i.e., diffusion coefficients, of the physical model are not known in advance and usually need to be identified. Generally, the problem of parameter identification leads to a nonlinear estimation problem. Hence, a novel efficient recursive procedure is employed. Unlike other estimators, the so-called *Hybrid Density Filter* not only assures accurate estimation results for nonlinear systems, but also offers an efficient processing. By this means it is possible to reconstruct and identify distributed phenomena monitored by autonomous wireless sensor networks. The performance of the proposed estimation method is demonstrated by means of simulations.

Keywords: Distributed phenomena, parameter identification, stochastic systems, nonlinear estimation, sensor-actuator-networks

I. INTRODUCTION

Miniaturization and recent developments in wireless sensor network technology make it possible to exploit sensor networks for monitoring natural large-area phenomena. The individual sensor nodes deployed in the environment provide distributed measurements in an intelligent and autonomous manner [1], [2] and collect information about the phenomenon to be monitored. Examples of distributed phenomena to be reconstructed include temperature distributions in a plane, fluid flows, deflection of bearings, and the surface motion of a beating heart in minimally invasive surgery [3].

For observing distributed phenomena by means of a sensor network, the individual sensor nodes are densely deployed either inside the phenomenon or very close to it. In such scenarios, the number of sensor nodes and the measurement rates should be as low as possible due to economical and energetic reasons. It can be stated that the lower the measurement rate of the sensor nodes, the higher their durability. Therefore, a trade-off between energy costs and accuracy has to be found. In addition, the challenge for the observation of distributed

phenomena is that measurements are available only at discrete time steps and discrete locations, meaning that no information between the individual sensor nodes are available. In order to get meaningful and accurate information not only at the sensor nodes itself but also between these nodes, the *model-based* reconstruction of the distributed phenomenon is of major significance [4]. By exploiting additional background information of the phenomenon in form of a physical model, the accuracy of the reconstruction can be improved significantly, even at non-measurement points. Accordingly, the number of sensor nodes and the measurement rates can be decreased.

The model-based reconstruction of a distributed phenomenon by means of a sensor network is based on an accurate physical model. Assuming we have an appropriate and sufficiently accurate physical model, the distributed phenomenon is uniquely characterized by model parameters and boundary conditions. However, in practical implementations the model parameters, such as diffusion coefficient and the boundary conditions of the distributed phenomenon might not be known in advance and usually need to be identified [5]. That means, by given sensor measurements, it is desirable to find the corresponding model parameters. By this means, the discrete-time samples measured by the individual sensor nodes are incorporated in the physical model in order to improve its accuracy in terms of estimated model parameters [6]. For sensor network applications, the parameter identification becomes even more essential due to the harsh and unknown environment, unpredictable variations of the phenomenon, and possibly unknown sensor locations.

In general, physical phenomena can be classified into *distributed*-parameter systems and *lumped*-parameter systems. The key characteristic of a lumped-parameter system is that the state, which uniquely describes the system behavior, depends only on time. In the case of a bird swarm or swarm of robots, the position of the individual robots, which can be regarded as the system state, depend only on time. Such a system can be conveniently described by a set of *ordinary* differential equations. On the other hand, the key characteristic of distributed-parameter systems is that the state not only depends on time but also on the location, e.g. irrotational

flow, heat conduction, and wave propagation. The behavior of distributed-parameter systems can be described by a set of *partial* differential equations.

The model-based reconstruction based on a distributed-parameter system description is quite complicated, because a system description in lumped-parameter form is typically required for standard Bayesian estimation approaches. That means, the physical model characterizing the distributed phenomena has to be converted from a distributed-parameter form to a lumped-parameter form. This conversion can be achieved by methods for solving partial differential equations, such as finite-difference method, finite-element method [7], modal analysis [8] and finite-spectral method [4], [9].

The novelty of this paper is the systematic approach to reconstruction and parameter identification for sensor network applications. Furthermore, the uncertainties occurring in the physical system and arising from noisy measurements are considered by an integrated treatment. By means of a model-based approach, it is possible to identify and track unpredictable variations both of the phenomena and of the sensor network itself. Hence, the properties and behavior of the sensor network can be autonomously adapted. Furthermore, the sensor network collects information about the phenomena to be monitored in terms of model parameters. The main challenge is that the simultaneous reconstruction and parameter identification of distributed phenomena automatically leads to nonlinear system and measurement models. Consequently a nonlinear estimation approach is required.

The Bayesian estimator provides a general framework for nonlinear system models affected by noise. However, exact Bayesian estimation is computationally demanding and a closed-form solution is impossible to obtain in general. Especially due to the limited resources in sensor networks, it is essential to work with an approximate estimator to keep the necessary computational demand bounded.

While widely-used approximate estimators like the Extended Kalman Filter (EKF) [10], Unscented Kalman Filter (UKF) [11], or even Particle Filters [12] require a trade-off between accuracy and efficiency, we utilize the Hybrid Density Filter (HDF) [13]. This estimator offers accurate estimation results due to a continuous non-Gaussian representation as well as an efficient processing due to a novel approximation technique based on a hybrid density.

The remainder of this paper is organized as follows: In Section II, a rigorous formulation of the problem and the main challenges of the simultaneous reconstruction and parameter identification of distributed phenomena is given. In Section III, the spatial and temporal decomposition is reviewed, which allows the approximation of the partial differential equation (distributed-parameter system) by means of a finite-dimensional system in state-space form (lumped-parameter system). The Section IV is devoted to the simultaneous reconstruction and parameter identification of distributed phenomena by means of a sensor network. It will turn out that the parameter identification usually leads to a nonlinear estimation problem, even in the simplest case. Accordingly, in

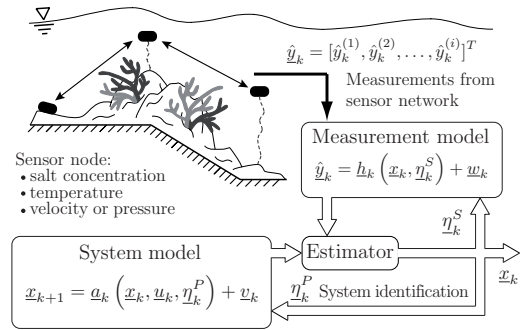


Fig. 1. Scenario for the reconstruction and the identification of a distributed phenomenon by means of a sensor network: Observation of a coral reef.

Section V a novel nonlinear filter, the so-called *Hybrid Density Filter*, is proposed. Finally, in Section VI, the performance of the proposed filter is demonstrated by means of simulation results: As an example, the temperature distribution in a heat conductor is reconstructed and the diffusion coefficient is estimated by incorporating measurements from a sensor network.

II. PROBLEM FORMULATION

The main goal is to design a dynamic lumped system model for the purpose of estimating the state of a distributed phenomenon monitored by a sensor network. Due to several unpredictable variations of the phenomenon and the sensor network itself, uncertainties naturally arise both in the dynamic system and from the measurements. The additional uncertainties in model parameters make the *simultaneous identification* essential for accurate reconstruction of distributed phenomena by means of a sensor network.

A large number of distributed phenomena, such as temperature distribution in a plate, heat conduction, flow in ducts and in open-channel, wave propagation, and deflection of bearings can be described by means of a set of linear partial differential equations. In this paper, only *one-dimensional linear partial differential equations* are considered for simplicity and brevity, although similar expressions can be found for the multi-dimensional nonlinear case.

In its most general form, the one-dimensional linear partial differential equation is given in implicit form by

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

where $p(z, t)$ denotes the state of the distributed phenomena, e.g. the temperature at a certain time t and certain location z . The source term driving the distributed phenomena is represented by $s(z, t)$. The system input $s(z, t)$, the state $p(z, t)$ and its derivatives are related by means of a linear operator denoted by $\mathbb{L}(\cdot)$. Furthermore, the behavior of the distributed phenomena strongly depends on specific, in general time-variant, physical parameters, such as diffusion coefficients or coefficients of viscosity. These parameters are denoted as $\eta^P(z, t)$ and generally depend on both location and time.

Please note that even in the simplest case, the unknown parameter vector $\underline{\eta}^P(z, t)$ is usually related to the system state $p(z, t)$ in a nonlinear way.

For solving the partial differential equation (1), and thus reconstructing the distributed phenomenon, knowledge of the boundary conditions is necessary. There are several types of boundary conditions depending on the constraints at the boundaries of the distributed phenomenon to be monitored. Considering the solution in a domain $\Omega = \{z|0 \leq z \leq L\}$, we assume the following boundary conditions

$$\frac{\partial p(z=0, t)}{\partial z} = g_N, \quad p(z=L, t) = g_D, \quad (2)$$

where g_N , specifying a condition on the derivative, is referred to as a *Neumann* boundary condition and g_D is the so-called *Dirichlet* boundary condition.

The aforementioned partial differential equation (1) can be regarded as the *infinite*-dimensional state-space form of the distributed phenomenon. However, the application of a Bayesian approach for reconstruction and identification purposes based on the *infinite*-dimensional system model is quite complex. For that reason, the partial differential equation (1) is approximated by means of a *finite*-dimensional system in state-space form.

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

$$\underline{x}_{k+1} = \underline{a}_k(\underline{x}_k, \underline{u}_k, \underline{\eta}_k^P) + \underline{w}_k^x, \quad (3)$$

where \underline{x}_k contains the individual states characterizing the time evolution of the distributed phenomenon, \underline{u}_k denotes the system input, and \underline{w}_k^x represents the system uncertainties.

In general, the parameter vector $\underline{\eta}_k^P$ in (3) 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 phenomenon, unknown system inputs, and unknown disturbances could be considered in the parameter vector $\underline{\eta}_k^P$ and therefore could be eventually identified. In the case of sensor-actuator-networks, the parameter vector $\underline{\eta}_k^P$ could even contain unknown properties of the actuators.

Besides the system model, which describes the physical behavior of the distributed phenomenon, there is a measurement model describing the physical properties of the sensor network itself. In this paper, we assume that the measurements \hat{y}_k are related *nonlinearly* to the state vector \underline{x}_k according to

$$\hat{y}_k = \underline{h}_k(\underline{x}_k, \underline{\eta}_k^S) + \underline{v}_k,$$

where \underline{v}_k is the uncertainty in the measurement model. The parameter vector $\underline{\eta}_k^S$ contains all the *unknown parameters* to be identified in the measurement model. Sensor bias and sensor variances, for example, could be included in the unknown parameter vector $\underline{\eta}_k^S$ for the purpose of tracking of physical

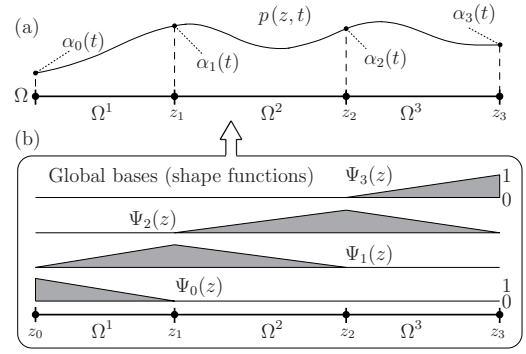


Fig. 2. The solution $p(z, t)$ of the partial differential equation is approximated by $\tilde{p}(z, t)$ which is characterized by shape functions $\Psi_j(z)$ and their weighting coefficients $\alpha_j(t)$. (a) Elemental decomposition of the solution domain Ω into several subdomains Ω^e . (b) Shape functions $\Psi_i(z)$ for the linear case.

wear of the sensor nodes. Furthermore, one could imagine to collect the possibly unknown location of the individual sensor nodes and correlations between the individual nodes in the parameter vector $\underline{\eta}_k^S$.

In this paper, only the identification of physical constants $\underline{\eta}_k^P$ in the system models, e.g. diffusion coefficient, are considered, although similar approaches could be derived for the identification of unknown measurement model parameters $\underline{\eta}_k^S$. It will be shown that the identification of the unknown parameter $\underline{\eta}_k^P$ leads to a nonlinear system model, and thus makes a nonlinear estimation approach necessary.

III. SPATIAL AND TEMPORAL DECOMPOSITION OF DISTRIBUTED SYSTEM

In this section, a method is introduced that allows the temporal and spatial decomposition of the partial differential equation, i.e., *distributed*-parameter system, leading to a *finite*-dimensional state-space form. As an example, we consider the following simple one-dimensional linear partial differential equation, the so-called diffusion equation,

$$\mathbb{L}(p(z, t)) = \frac{\partial p(z, t)}{\partial t} - \eta^P(t) \frac{\partial^2 p(z, t)}{\partial z^2} - s(z, t) = 0, \quad (4)$$

with the boundary condition stated in (2) and the unknown diffusion coefficient $\eta^P(t)$ that needs to be identified by incorporating sensor measurements, see Sec. IV-B.

A. Spatial Decomposition

By means of spatial decomposition, it is possible to approximate the partial differential equation (4) by a set of ordinary differential equations. For that purpose, the solution domain $\Omega = \{z|0 \leq z \leq L\}$ needs to be decomposed into N_{el} subdomains Ω^e the so-called *finite elements*. These subdomains are defined by the element boundary nodes z_i for $i = 0, 1, \dots, N_{dof} - 1$, where $z_0 = 0$ and $z_{N_{dof}-1} = L$. The elemental decomposition of the solution domain Ω into several subdomains Ω^e is visualized by means of an example in Fig. 2.

It is well-known that the finite-difference, the finite-element, and the finite-spectral method may be used with the same numerical methodology, the so-called *Galerkin* formulation. For this method, it is assumed that the solution $p(z, t)$ in the solution domain Ω can be represented by a piecewise approximation $\tilde{p}(z, t)$ according to

$$\tilde{p}(z, t) = \sum_{i=0}^{N_{\text{dof}}-1} \psi_i(z) \alpha_i(t) = \underline{\Psi}^T(z) \underline{x}(t) , \quad (5)$$

where $\Psi_i(z)$ are analytic functions called the *shape functions*. It is important to note that the individual shape functions $\Psi_i(z)$ are defined in the entire solution domain Ω and the form strongly depends upon the used numerical method, e.g., orthogonal polynomials for finite-spectral method [4].

Due to the fact that the approximated solution $\tilde{p}(z, t)$ cannot satisfy the partial differential equation (4) everywhere in the region of interest, a *residual* R_Ω remains. To make this residual small in some sense, a *weighted integral* has to be minimized

$$\int_{\Omega} \Psi_i(z) \mathbb{L}(p(z, t)) dz = 0 ,$$

with $i = 0, 1, \dots, N_{\text{dof}} - 1$.

By replacing both the solution function $p(z, t)$ and the input function $s(z, t)$ by the finite expression (5), the weighted integral can be reduced to a system of *ordinary differential equations* in terms of the weighting coefficients $\alpha_i(t)$,

$$\mathbf{M}_G \dot{\underline{x}}(t) = -\eta^P(t) \mathbf{D}_G \underline{x}(t) + \eta^P(t) \underbrace{(\mathbf{M}_G \underline{u}^*(t) + \underline{b}^*(t))}_{\underline{u}(t)} , \quad (6)$$

where \mathbf{M}_G is called the global *mass* matrix and \mathbf{D}_G is the global *diffusion* matrix. The individual entries M_{ij}^g and D_{ij}^g , which assemble to the global mass matrix \mathbf{M}_G and the global diffusion matrix \mathbf{D}_G , can be derived according to

$$M_{ij}^g = \int_{\Omega} \Psi_i(z) \Psi_j(z) dz , \quad D_{ij}^g = \int_{\Omega} \frac{d\Psi_i(z)}{dz} \frac{d\Psi_j(z)}{dz} dz .$$

It is obvious that \mathbf{M}_G and \mathbf{D}_G contain the information about the discretized domain Ω and merely depend upon the choice of the shape functions $\Psi_i(z)$, i.e., depend on the used approximation method. The vectors $\underline{x}(t)$ and $\dot{\underline{x}}(t)$ are the state vectors of the unknown weighting coefficients $\alpha_i(z)$ and their derivatives

$$\underline{x}(t) = [\alpha_0(t), \alpha_1(t), \dots, \alpha_{N_{\text{dof}}-1}(t)] .$$

The boundary conditions of the distributed phenomenon to be monitored are contained in the boundary condition vector $\underline{b}^*(t)$. For brevity, the input vector $\underline{u}^*(t)$ and the boundary condition vector $\underline{b}^*(t)$ can be combined to a so-called augmented input vector $\underline{u}(t)$. The interested reader should refer to [4] and [9] for more information on how to derive the system of ordinary differential equations (6).

B. Temporal Decomposition

In the previous section, we presented the spatial decomposition allowing the conversion of the partial differential equation into a set of ordinary differential equations, i.e., the conversion of the distributed-parameter system into a lumped-parameter system. In this section we are now ready to specify the time evolution leading to the discrete-time system model.

The time discretization of the set of ordinary differential equations (6) by means of an *explicit* discretization method leads to following equation

$$\mathbf{M}_G \frac{\underline{x}_{k+1} - \underline{x}_k}{\Delta t} = -\eta_k^P \mathbf{D}_G \underline{x}_k + \eta_k^P \underline{u}_k ,$$

where Δt represents the time step and has to be chosen in such a way that the resulting system model is *numerically stable*. Rearranging finally leads to the lumped-parameter system

$$\underline{x}_{k+1} = \mathbf{A}_k(\eta_k^P) \underline{x}_k + \mathbf{B}_k(\eta_k^P) \underline{u}_k , \quad (7)$$

where the system matrix \mathbf{A}_k and the input matrix \mathbf{B}_k depend on the unknown parameter η_k^P to be identified. These matrices can be derived according to

$$\begin{aligned} \mathbf{A}_k(\eta_k^P) &= \mathbf{I} - \Delta t \mathbf{M}_G^{-1} \mathbf{D}_G \eta_k^P , \\ \mathbf{B}_k(\eta_k^P) &= \Delta t \mathbf{M}_G^{-1} \eta_k^P . \end{aligned}$$

Please note that we have chosen an explicit time discretization method for simplicity and brevity. To circumvent the restriction on the time step Δt , it is reasonable to integrate the set of ordinary differential equations by means of *implicit* methods, such as the Crank-Nicolson discretization as it was shown in [4]. These methods lead to a system model which is *unconditionally* stable for any time step Δt .

The *finite*-dimensional system model (7) represents the approximation of the *infinite*-dimensional state-space form of the distributed phenomenon characterized by the partial differential equation (4). In the next section, the finite-dimensional system model is used for the reconstruction and parameter identification of the distributed phenomena.

IV. SIMULTANEOUS RECONSTRUCTION AND PARAMETER IDENTIFICATION

The physical model in finite-dimensional state-space form (7) derived in the last section could be used for the simulation of the underlying distributed phenomenon. However, for applications in sensor networks the aim is not only to simulate the behavior of the phenomenon but also to reconstruct and identify the distributed phenomenon by incorporating discrete-time measurements. The individual sensor nodes, which are densely deployed inside the distributed phenomenon, are exploited to monitor and identify that phenomenon and thus to improve both the estimated state \underline{x}_k and the unknown parameter vectors $\underline{\eta}_k^P$ and $\underline{\eta}_k^S$.

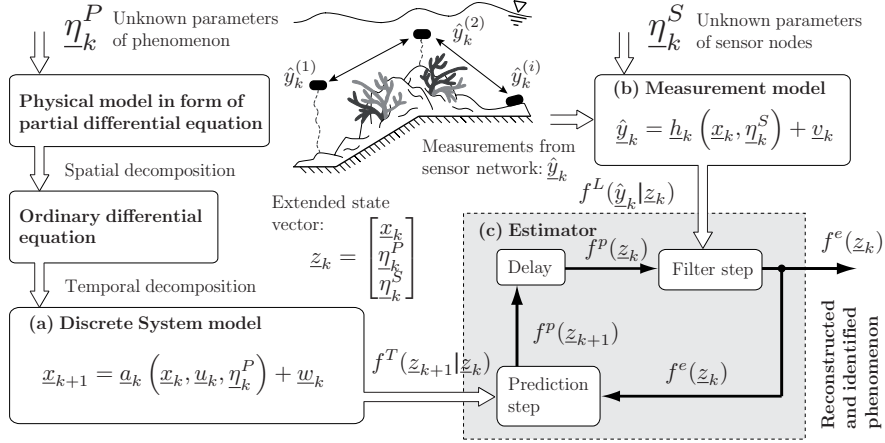


Fig. 3. Overview and components of the procedure for model-based simultaneous parameter identification and reconstruction of distributed phenomena. (a) System model, (b) Measurement model, and (c) Estimator.

A. Model-based Reconstruction of Distributed Systems

In addition to a system model a measurement model, which maps the specific measurements \hat{y}_k obtained from the sensor network to the solution $p(z, t)$ of the observed phenomenon is necessary. The measurement model consists of two parts: Measurement equation and output equation.

The measurement equation relates the actual measurement $\hat{y}_k^{(i)}$ at location $z^{(i)}$ and at time t_k to the distributed phenomenon $p(z, t_k)$ according to

$$\hat{y}_k = \underline{h}_k^* \left(p(z, t_k), \eta_k^S \right) + v_k ,$$

where the measurement vector \hat{y}_k contains the actual measurements of the m sensor nodes, i.e., $\hat{y}_k = [\hat{y}_k^{(1)}, \dots, \hat{y}_k^{(m)}]^T$. The vector v_k represents the uncertainties arising from the actual sensor node and the parameter vector η_k^S contains the imprecisely known sensor node properties, e.g. sensor bias, sensor variances or sensor locations.

The output equation, on the other hand, relates the measurements of the distributed phenomenon $p(z^{(i)}, t_k)$ directly to the finite-dimensional state vector \underline{x}_k , according to

$$p(z^{(i)}, t_k) = \sum_{j=0}^{N_{\text{dof}}-1} \psi_j(z^{(i)}) x_k^{(j)} ,$$

which is identical to the representation of the approximate solution $\tilde{p}(z^{(i)}, t_k)$ for the partial differential equation.

Substitution of the output equation into the measurement equation leads to the complete measurement model

$$\hat{y}_k = \underline{h}_k \left(\underline{x}_k, \eta_k^S \right) + v_k .$$

This equation provides the mapping of the finite-dimensional state vector \underline{x}_k to the individual discrete-time measurements \hat{y}_k obtained from the sensor network.

Assuming a linear measurement equation for a sensor network consisting of L sensor nodes, the measurement gain

matrix \mathbf{H}_k is set up by the shape function $\underline{\Psi}^T(z^{(i)})$ of order N , leading to the following linear measurement model

$$\hat{y}_k = \underbrace{\begin{bmatrix} \psi_1(z^{(1)}) & \dots & \psi_N(z^{(1)}) \\ \vdots & \ddots & \vdots \\ \psi_1(z^{(L)}) & \dots & \psi_N(z^{(L)}) \end{bmatrix}}_{\mathbf{H}_k} \underline{x}_k + v_k ,$$

where v_k are the measurement uncertainties; see [4] for more information.

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 reconstruct the entire phenomenon.

B. Simultaneous Reconstruction and Parameter Identification

In the last section, it was assumed that the parameter vectors η_k^P and η_k^S are precisely known, so that the distributed phenomenon can be reconstructed by means of the sensor network. In this section, we assume that the parameter vectors η_k^P and η_k^S are not initially known, and thus need to be identified. The parameters characterizing the distributed phenomenon and the sensor network are usually time-variant and affected by disturbances not considered in the respective models. This criterion becomes much more essential in a sensor network application due to many uncertainties, such as on node locations, on sensor properties, on boundary conditions, and on physical properties of the phenomena itself. Consequently, we are proposing to reconstruct the distributed phenomenon and identify the unknown parameters simultaneously, as illustrated in Fig. 3.

For simultaneous reconstruction and parameter identification, the unknown parameter vectors η_k^P and η_k^S are treated as additional state variables. Hence, an *augmented state vector* \underline{z}_k containing the system state \underline{x}_k and the additional unknown

parameters can be defined according to

$$\underline{z}_k := \begin{bmatrix} \underline{x}_k^P \\ \eta_k^P \\ \eta_k^S \end{bmatrix}.$$

Assuming the unknown parameters η_k^P and η_k^S to be slowly time varying we can model these parameters as follows

$$\eta_{k+1}^P = \eta_k^P + \underline{w}_k^P, \quad \eta_{k+1}^S = \eta_k^S + \underline{w}_k^S,$$

where \underline{w}_k^P and \underline{w}_k^S are small artificial noise terms that allow the estimator to modify the estimate of η_k^P and η_k^S respectively.

The augmentation of the state vector with additional unknown parameters leads to the so-called *augmented system model*. In the case of simultaneous reconstruction and identification of the diffusion coefficient of the diffusion equation (4), the augmentation leads to the following augmented system model

$$\underbrace{\begin{bmatrix} \underline{x}_{k+1}^P \\ \eta_{k+1}^P \\ \eta_{k+1}^S \end{bmatrix}}_{=\underline{z}_{k+1}} = \underbrace{\begin{bmatrix} \mathbf{A}_k(\eta_k^P)\underline{x}_k + \mathbf{B}_k(\eta_k^P)\underline{u}_k \\ \eta_k^P \\ \eta_k^S \end{bmatrix}}_{=a_k(\underline{z}_k)} + \underbrace{\begin{bmatrix} \underline{w}_k^P \\ \underline{w}_k^S \end{bmatrix}}_{=\underline{w}_k} \quad (8)$$

and measurement model

$$\hat{y}_k = \underline{h}_k(\underline{x}_k, \eta_k^S) + v_k,$$

where \underline{w}_k and v_k denotes the system noise and the measurement noise respectively. In this case, it is obvious that the augmented system model is *nonlinear* in the augmented state \underline{z}_k due to the multiplication of $\mathbf{A}_k(\eta_k^P)$ containing the unknown parameter η_k^P and the system state \underline{x}_k .

V. HYBRID DENSITY FILTER (HDF)

For nonlinear systems that are affected by noise, the exact Bayesian estimator calculates the probability density function representing the augmented state. It consists of the so-called *prediction step* for a recursive state propagation in time and a *measurement update* or filter step to incorporate actual measurements (see Fig. 3 (c)). However, due to its high computational demand and the resulting non-parametric density representation, an approximate estimation is inevitable as mentioned in Section I. The Hybrid Density Filter (HDF) provides the opportunity of accurate estimations in combination with an efficient processing due to its novel approximation technique [13].

A. Prediction Step

In this section, a short review of the prediction step of the HDF is provided. Given the current probability density function $f_k^e(\underline{z}_k)$ of the augmented state at time step k , the Chapman-Kolmogorov equation [14]

$$f_{k+1}^p(\underline{z}_{k+1}) = \int_{\mathbb{R}^n} f^T(\underline{z}_{k+1}|\underline{z}_k) f_k^e(\underline{z}_k) d\underline{z}_k \quad (9)$$

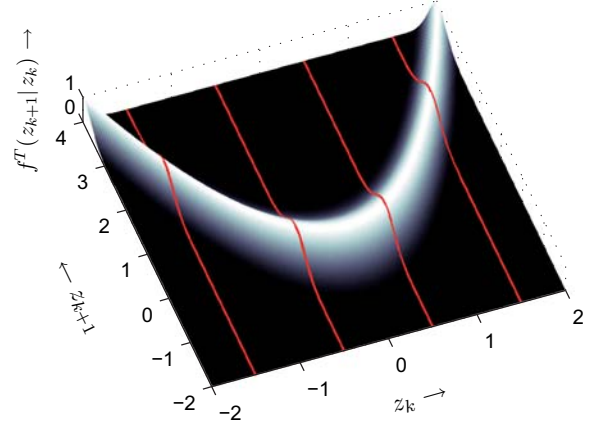


Fig. 4. The hybrid transition density approximation with $K = 4$ components (red lines) slices the true transition density in 5 parts. Each slice is a Gaussian density.

describes the prediction of the Bayesian estimator and derives the *predicted density* $f_{k+1}^p(\underline{z}_{k+1})$, i.e., the density of \underline{z}_k at the next time step $k + 1$. In (9), $f^T(\underline{z}_{k+1}|\underline{z}_k)$ given by

$$f^T(\underline{z}_{k+1}|\underline{z}_k) = f_k^w(\underline{z}_{k+1} - a_k(\underline{z}_k))$$

is the so-called *transition density*, which strongly relies on the nonlinear augmented system model (8) and the noise \underline{w}_k affecting it. For simplicity, we assume that this noise is zero-mean white Gaussian with density

$$\underline{w}_k \sim f_k^w(\underline{w}_k) = \mathcal{N}(\underline{w}_k - \underline{\mu}_k^w, \mathbf{C}_k^w),$$

where $\underline{\mu}_k^w = \underline{0}$ is the mean vector and \mathbf{C}_k^w is the covariance matrix. However, the HDF is not restricted to Gaussian noise.

While most of the state-of-the-art estimators concentrate on directly approximating the density of the state, the key idea of the HDF is to approximate the transition density instead, which is the probabilistic representation of the nonlinear system model. For approximating the transition density a *hybrid density*

$$f(\underline{z}_{k+1}, \underline{z}_k, \underline{\kappa}) = \sum_{i=1}^K \omega_i \cdot \delta(\underline{z}_k - \underline{\mu}_i^{(1)}) \cdot \mathcal{N}(\underline{z}_{k+1} - \underline{\mu}_i^{(2)}, \mathbf{C}_i) \quad (10)$$

with parameter vector

$$\underline{\kappa} = [\underline{\kappa}_1^T, \dots, \underline{\kappa}_K^T]^T, \quad \text{where } \underline{\kappa}_i^T = [\omega_i, \underline{\mu}_i^{(1)}, \underline{\mu}_i^{(2)}, \mathbf{C}_i]^T$$

is used. Here, K is the number of components, ω_i are weight coefficients with $\omega_i > 0$, and $\delta(\underline{z}_k - \underline{\mu}_i^{(1)})$ is the Dirac delta function with mean vector $\underline{\mu}_i^{(1)} = [\mu_{i,1}^{(1)}, \dots, \mu_{i,n}^{(1)}]^T$. Hence, the hybrid density is a mixture density consisting of Dirac components for the prior side \underline{z}_k and Gaussian components for the posterior side \underline{z}_{k+1} .

To adjust these parameters, it is possible to reformulate the approximation problem as an optimization problem that yields the optimal parameter vector $\underline{\kappa}$ with respect to an appropriate

distance measure [13]. Assuming that $\underline{z}_k \in S_k$, i.e., \underline{z}_k is restricted to its support

$$S_k = [s_1^l, s_1^u] \times [s_2^l, s_2^u] \times \cdots \times [s_n^l, s_n^u] \subset \mathbb{R}^n,$$

such that $f_k^e(\underline{z}_k) > \epsilon$ for $\underline{z}_k \in S_k$ and for a small constant $0 < \epsilon \ll 1$, the optimal parameters are given by

$$\begin{aligned} \omega_i &= \prod_{j=1}^n \omega_{i,j}, & \omega_{i,j} &= \frac{s_j^u - s_j^l}{L_j}, \\ \mu_{i,j}^{(1)} &= s_j^l + \omega_{i,j} \cdot \frac{2i-1}{2}, & \mu_i^{(2)} &= a_k(\mu_i^{(1)}), \\ \mathbf{C}_i &= \mathbf{C}_k^w, \end{aligned}$$

for $j = 1, \dots, n$. In Fig. 4 the transition density for the scalar nonlinear system model $z_{k+1} = z_k^2 + w_k$, where $w_k \sim \mathcal{N}(w_k - 0, 0.16)$, is depicted. The corresponding hybrid density approximation consists of $K = 4$ components, whose Dirac delta functions are placed uniformly on the interval $S_k = [-2, 2]$. The Gaussian elements are duplicates of the noise density that are placed along the nonlinear function $a_k(z_k) = z_k^2$.

By using (10) instead of the transition density in (9), prediction can be performed efficiently in closed-form, resulting in a Gaussian mixture representation

$$f_{k+1}^p(\underline{z}_{k+1}) = \sum_{i=1}^K \omega_{k+1,i} \cdot \mathcal{N}(\underline{z}_{k+1} - \mu_i^{(2)}, \mathbf{C}_i), \quad (11)$$

with $\omega_{k+1,i} = \omega_i \cdot f_k^e(\mu_i^{(2)})$, for the predicted density. Finally, normalizing (11) by multiplying it with $1/\sum_{i=1}^K \omega_{k+1,i}$ yields a very accurate approximation of the true predicted density [13].

B. Measurement Update

In the same way as the prediction step can be performed by means of hybrid densities, we can use a hybrid density approximation of the likelihood $f^L(\hat{y}_k | \underline{z}_k)$ in Bayes' law [14]

$$f_k^e(\underline{z}_k) = c_k f^L(\hat{y}_k | \underline{z}_k) f_k^p(\underline{z}_k) \quad (12)$$

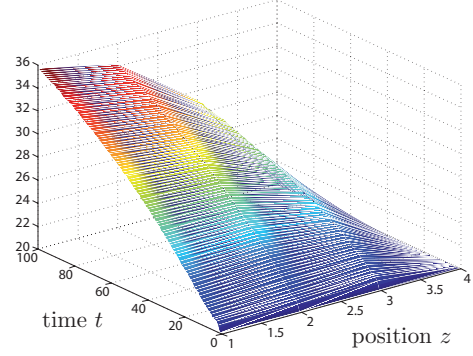
for updating the augmented state \underline{z}_k with an actual measurement \hat{y}_k to achieve the posterior density $f_k^e(\underline{z}_k)$. This approach is described in detail in [13].

In this paper, we use a more straightforward approach to further decrease the computational load. Instead of a separate measurement update step, we propose a combined prediction and update by plugging (12) into (9). Consequently, we obtain

$$f_{k+1}^p(\underline{z}_{k+1}) = c_k \int_{\mathbb{R}^n} f^T(\underline{z}_{k+1} | \underline{z}_k) f^L(\hat{y}_k | \underline{z}_k) f_k^p(\underline{z}_k) d\underline{z}_k,$$

where $c_k = 1/\int_{\mathbb{R}^n} f^L(\hat{y}_k | \underline{z}_k) f_k^p(\underline{z}_k) d\underline{z}_k$ is merely a normalization constant. Here, it is sufficient to approximate the transition density only and to perform the prediction step as described before for additionally updating the system state by means of an actual measurement. According to (11), a Gaussian mixture representation for $f_{k+1}^p(\underline{z}_{k+1})$ is obtained. However, by performing a combined update and prediction it is impossible to achieve a separate posterior density $f_k^e(\underline{z}_k)$, which is not always preferable.

(a) Estimated solution



(b) Elemental decomposition

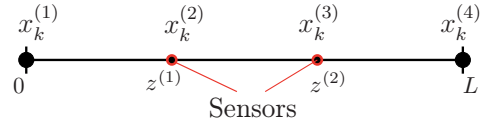


Fig. 5. (a) Estimated solution $\tilde{p}(z, t)$ for precisely known measurement model parameter η_k^S and system model parameter η_k^P , i.e., $\eta_k^P = 0.6$ (b) Elemental decomposition and location of sensor nodes.

VI. SIMULATION RESULTS

In this section, the performance of the proposed estimation method is demonstrated by means of simulations for the reconstruction and identification of a heat rod. In particular, the accuracy of the reconstructed phenomenon and of the identified parameter η_k^P characterizing the partial differential equation is investigated in comparison to other nonlinear estimation methods.

The evolution of the temperature in the heat rod with length $L = 1$ is modeled by the one-dimensional linear partial differential equation (4) with initial condition $p(z, t = 0) = 20$, Dirichlet boundary condition $p(z = L, t) = 20$, and Neuman boundary condition $\frac{\partial p(z=0, t)}{\partial t} = 0$. The noisy input function is given by $s(z, t) = 50$. Approximating the solution $p(z, t)$ by means of $\tilde{p}(z, t) = \underline{\Psi}^T(z) \underline{\alpha}(t)$, the partial differential equation (6) can be spatially decomposed leading to the *finite*-dimensional state-space form (7). The state vector \underline{x}_k can be derived from temporal discretization of the weighting factors $\underline{\alpha}(t)$ of the approximate solution, as shown in Sec. III. For simplicity, the shape functions $\psi_i(z)$ are assumed to be piecewise linear. Furthermore, it is assumed that the sensor network consists of two sensor nodes located at $z^{(1)} = 0.33$ and at $z^{(2)} = 0.66$, as it is visualized in Fig. 5. As a measurement model we assume $\hat{y}_k = \mathbf{I} \underline{x}_k + \underline{v}_k$, where \mathbf{I} is the identity.

For simulation we investigate the Unscented Kalman Filter (UKF), the Particle Filter (PF), and the proposed HDF for identifying the parameter η_k^P , which is assumed to have the value 0.6. PF and HDF operate with 405 particles and components respectively. In Fig. 6 (a) and (b) the mean estimate of the identified parameter η_k^P and the absolute difference to the true parameter are visualized. It is obvious

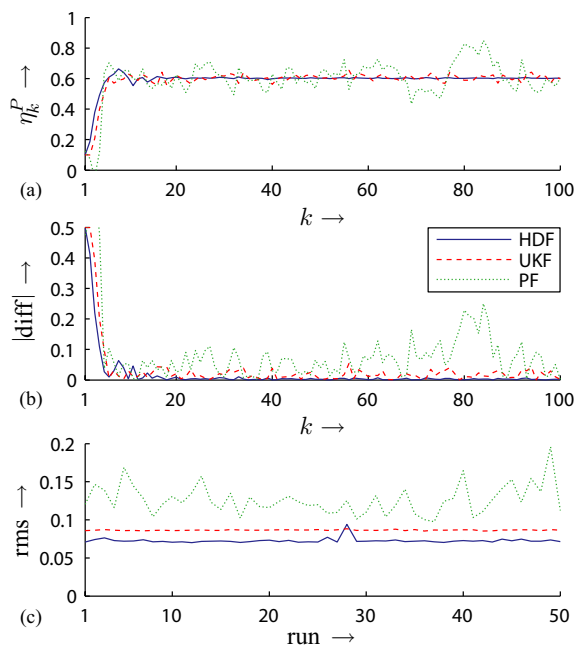


Fig. 6. Comparison of the HDF, UKF, and PF for $\eta_k^P = 0.6$. (a) Mean estimation and (b) absolute difference for an exemplary simulation run. (c) Root mean square error of 50 simulation runs.

that the HDF offers parameter estimation, which is nearly exact, since the hybrid transition density approximation allows to cover the nonlinearity of the system model. Therefore, a more accurate estimation of moments and shape of the predicted density is provided. The UKF provides also a good parameter identification result, while the PF strongly jitters. Increasing the number of particles in turn results in a higher computational load.

The estimates depicted in Fig. 6 (a) and (b) are resulting from one simulation run out of 50 Monte Carlo simulations. The root mean square errors (rms) of all 50 runs are depicted in Fig. 6 (c). Except of one simulation run, the HDF always outperforms the UKF and PF. Thanks to the model-based approach the unknown parameter η_k^P can be identified and the solution $p(z, t)$ in the entire domain can be reconstructed by incorporating measurements from the sensor network.

VII. CONCLUSIONS AND FUTURE WORK

This paper introduced the methodology for deriving system models and measurement models for the reconstruction and identification of distributed phenomena characterized by linear partial differential equations. By means of the augmentation of the system state with the unknown parameters it was possible to simultaneously estimate the system state and identify the unknown parameter. It was shown that the augmentation of the state vector leads to nonlinear system models. In order to cope with these nonlinearities, a novel estimator, the so-called *Hybrid Density Filter*, was proposed and its performance was demonstrated by means of simulations. It turned out that, compared to other nonlinear estimators, the *Hybrid Density Filter* achieves a higher accuracy.

The application of such reconstruction and identification methods to sensor network applications provides novel prospects. The individual sensor nodes would be able to operate in a completely unknown surrounding in a self-organized and autonomous manner and thereby identify, observe, and reconstruct unknown distributed phenomena.

For considering the physical wear of individual nodes in a sensor network the identification of sensor properties such as sensor bias, measurement variance, and node location is of special interest. This is left for future research work. Furthermore, for a lot of real sensor network applications it might be necessary to identify complex, not a-priori known, boundary conditions. For the observation of large-area distributed phenomena, decentralized methods or modularization methods could be necessary in order to cope with high-dimensional state vectors.

VIII. 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] B. Thuraisingham, “Secure sensor information management and mining,” *IEEE Signal Processing Magazine*, vol. 5, pp. 14–19, 2004.
- [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] 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)*, Setubal, Portugal, August 2006, pp. 16 – 23.
- [5] 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.
- [6] 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.
- [7] A. J. Baker, *Finite Element Computational Fluid Mechanics*. Taylor and Francis, 1983.
- [8] K. Roberts and U. D. Hanebeck, “Prediction and Reconstruction of Distributed Dynamic Phenomena Characterized by Linear Partial Differential Equations,” in *The Eighth International Conference on Information Fusion (FUSION)*, 2005.
- [9] G. E. Karniadakis and S. Sherwin, *Spectral/hp Element Methods for Computational Fluid Dynamics*. Oxford University Press, 2005.
- [10] H. W. Sorenson, *Kalman Filtering: Theory and Application*. Piscataway, NJ: IEEE, 1985.
- [11] S. J. Julier and J. K. Uhlmann, “Unscented Filtering and Nonlinear Estimation,” in *Proceedings of the IEEE*, vol. 92, no. 3, 2004, pp. 401–422.
- [12] M. S. Arulampalam, S. Maskell, N. Gordon, and T. Clapp, “A Tutorial on Particle Filters for Online Nonlinear/Non-Gaussian Bayesian Tracking,” *IEEE Transactions on Signal Processing*, vol. 50, no. 2, pp. 174–188, February 2002.
- [13] M. F. Huber and U. D. Hanebeck, “The Hybrid Density Filter for Nonlinear Estimation based on Hybrid Conditional Density Approximation,” in *The Tenth International Conference on Information Fusion (FUSION)*, 2007.
- [14] F. C. Schweppe, *Uncertain Dynamic Systems*. Prentice-Hall, 1973.