Distributed Parameter Estimation for Monitoring Diffusion Phenomena Using Physical Models

Lorenzo A. Rossi, Bhaskar Krishnamachari and C.-C. Jay Kuo
Department of Electrical Engineering, University of Southern California
Los Angeles, CA 90089-2564, USA.
E-mails: lrossi@usc.edu, bkrishna@usc.edu, cckuo@sipi.usc.edu

Abstract— **In this work we address the problem of estimating parameters of diffusion phenomena via autonomous wireless sensor networks. Diffusion phenomena, such as the propagation of a gas in the air or of a chemical agent in the water, can be modeled by means of partial differential equations (PDE's). In several scenarios, the parameters characterizing such models, i.e. the coefficients of the PDE's, are not known a-priori and need to be estimated. We develop an adaptive approach for the distributed identification of the parameters of diffusion models for both the cases of known and unknown boundary conditions (BC's). The technique also applies to the case of spatially varying parameters. We present simulation results to show the performance and the various trade-offs of the method.**

Keywords— **Sensor networks, physical models, partial differential equations, diffusion phenomena, distributed identification, prediction, data gathering.**

I. Introduction

Wireless sensor networks aim to provide a smart interaction with the physical world. They can be deployed at low cost and in large numbers in remote environments to provide autonomous and intelligent measurements, answer queries, perform monitoring tasks. An important set of application scenarios consists in monitoring phenomena having a diffuse extension over the space. Examples of relevant phenomena include the temperature field in an ecosystem, clouds of gases in the air, polluting agents in the sea. In such cases, a sensor network may be required to transmit and keep updated a map of the phenomenon to a remote base station (BS) and to answer queries from the BS.

A widely studied class of approaches to handle the above problems consists simply in continuously transferring all the samples of the phenomenon taken at the various sensor locations to the BS. This is often referred as data gathering. In this contest people study the optimal way (in the sense of least energy consumption) to transfer node measurements to a remote base station (BS). A typical assumption is that all nodes sample the sensor field uniformly in time and generate a packet for each round of measurements. All packets are delivered to the BS through various aggregation strategies [1].

It must be remarked that many of the the aforementioned methods implicitly treat the sampled data as if they were an i.i.d. process. However, natural phenomena are typically characterized by strong spatial and temporal correlations and therefore the i.i.d assumption may lead to methods redundant in computation and communication. Recall that usually sensor network devices are severely energy limited. In many real world applications, such correlations can be characterized by relatively simple physical models, *e.g.* one or a couple of partial differential equations (PDE's). For instance, partial differential (diffusion) equations can be used to model phenomena such as the propagation of a gas in the air and of a polluting fluid in the water, or the evolution of the temperature in a material [2].

Therefore, physical models can give fundamental insights on many phenomena potentially interesting to sensor network applications. If this is the case, exploiting physical models in designing wireless sensor networks algorithms can lead to very appealing applications as it embeds useful a-priori knowledge of the phenomenon. For example, physics can be used to design prediction schemes or to provide a compact parametric description of the phenomenon [3], or to design more efficient encoding schemes for sensor data [4].

A relevant problem in the aforementioned framework is the identification of unknown parameters (e.g. the diffusion coefficient of a certain gas) of physical models through sample sensor data. The knowledge of the parameters, along with boundary and initial conditions, allows the complete characterization of the monitored phenomena and can be potentially exploited in many sensor network applications. There are many mathematical challenges in the estimation of the parameters of diffusion models even in a traditional centralized scenario [5]. Those challenges become much harsher in a sensor network scenario, due to the well known limitations in computation and communication of those devices, but also due many uncertainties (e.g. on node locations, on boundary conditions and sources) likely to be associated to those environments.

In this work, we assume the existence of an underlying diffusion PDE model for the physical phenomenon being monitored by the sensor networks (e.g. propagating gasses). To our knowledge, this research is one of the first studies on distributed identification of PDE parameters. In particular, we study the problem under different kinds of boundary conditions and with a-priori knowledge of the boundaries as well without it. The latter can be the case of several real scenarios where the sensor network must estimate initially unknown information about boundaries. Besides, the boundary conditions can be used to model phenomena affected by sources external to the sensor field. Furthermore, we consider also the case of parameters vary-

ing with the space.

The rest of this paper is organized as follows. Section II gives a summary over related works on non-sensor network centralized approaches for the identification of coefficients of diffusion models. The PDE diffusion model is introduced in Section III. The problem we address is presented in Section IV. Sections V and VI present our approach to the estimation of the parameters for the cases of known and unknown boundaries respectively. Section VII discusses the issues related to the relaxation of some assumptions to more realistic scenarios. Finally simulation results are shown and discussed in Section VIII.

II. Review of Related Work

The assumption of i.i.d. processes, which is implicit to many data gathering approaches [1], is in general unrealistic for modeling physical phenomena sampled by a sensor network. In many scenarios, there do exist some correlation of measured data, which can be exploited to develop a more effective data gathering system. One such example is the monitoring of the diffusion of some substance (e.g. the propagation of toxic gas). The sensor network sends to the BS the spatio-temporal samples of the phenomenon of interest, which have strong correlations. The exploitation of such correlations by means of a physical model, which corresponds to the diffusion PDE in this research, has the potential to reduce the amount of data to be transfered to the BS. There is not yet much research on physical modeling problems in sensor networks. Two very recent works in such areas are [3, 4]. In [4] some encoding schemes exploiting the heat diffusion model are derived.

A key problem in physics based sensor networks is the identification of model parameters (the diffusion coefficients) via in-network distributed processing. Centralized algorithms for the identification of PDE coefficients have been presented in the literature (see the references in the next paragraph). However, to the best of our knowledge, there is little research on a distributed approach to achieve this goal, except for our preliminary work [3].

The problem of identification of parameters of diffusion models has been investigated throughout mathematical and numerical methods by people in the area of material sciences [6, 7]. Besides, the identification and control of systems modeled by PDE's have been studied by researchers in the fields automatic control and applied mathematics [5, 8]. In this context, systems represented by PDE's are referred to as *infinite dimensional* or *distributed* systems[1] [9]. Moreover, a *lumped system* indicates a system that has discrete components (e.g. the discretized version of an infinite dimensional system). Here, we follow the terminology of system theory, but reserve the term *distributed algorithms* to indicate algorithms using multi-point independent computations and term *distributed computation* with the same meaning as commonly used in the sensor network context.

[1]Notice that the term *distributed systems* has a different meaning from those normally given in the sensor network literature.

Several adaptive algorithms for the identification of parameters of infinite dimensional systems in the continuous time-space domain have been presented [5, 8]. One major application of this study is the automatic control of PDE-modeled industrial plants. Adaptive control schemes that rely on the estimation system states and unknown parameters of a plant were proposed by Baumeister *et al.* [8]. They also developed a theory on the use of a finite dimensional approximation to the infinite dimensional estimator. Orlov and Bentsman [5] studied the problem of identifying an infinite dimensional system with spatial-varying parameters. They derived some constructive identifiability conditions that allow to determine persistently exciting input sources. In [10], Syrmos *et al.* formulated the problem via discrete-time state space equations and then estimated the parameters using nonlinear filtering techniques such as the of the extended Kalman filter (EKF).

All aforementioned algorithms are centralized and sensors are assumed to be connected to the computing center directly. Besides, the computing center is assumed to have perfect knowledge of boundary conditions and manipulability of sources. These assumptions are reasonable in classic system control and identification problems, but not applicable in a wireless sensor network environment. In the latter context, the amount of data being exchanged among nodes is a critical parameter due to the limited amount of energy in sensor nodes and the relatively high communication cost. We address some of the above constraints in this work, as described in the next sections.

III. The Diffusion Model

This section introduces the physical model that we adopt in this research. Consider a physical phenomenon represented by the space and time varying scalar field $x(\xi, t)$, where $0 < \xi < L$ and $t > 0$. If $x(\xi, t)$ follows a diffusion model, then the following partial differential equation holds [2]:

$$x_t(\xi, t) = D x_{\xi\xi}(\xi, t), \tag{1}$$

where the scalar D is called *diffusivity* and $x_t(.)$ and $x_{\xi\xi}(.)$ denote respectively the first order time derivative and the second order spatial derivative of $x(.)$. The above parabolic PDE is called *diffusion equation.*

Typically $x(\xi, t)$ represents a concentration. The meaning of equation (1) is that there is net flow of substance from the regions with higher concentration of the substance to the ones of lower concentration. Heat diffusion is another example of diffusion process. In such a case eq. (1) is called the *heat equation* and $x(.)$ represents a temperature, $T(.)$.[2] An interesting property of the diffusion equation is the *smoothing* effect: thanks to the proportionality between time derivative $x_t(\xi, t)$ and local curvature $x_{\xi\xi}(\xi, t)$, eq. (1) will have over the time a low pass filtering action on the spatial profile of $x(\xi, t)$.

[2]For simplicity we consider a mono-dimensional (1-D) case in space. The extension to a higher dimensional case is straightforward [3]. However, note that equation (1) can already model real phenomena such as the diffusion of fluid in a water channel or the propagation of heat in a metallic rod.

The diffusion equation can be solved with the knowledge of the initial condition (IC), i.e. $x(0,\xi) = x_0(\xi)$ and the boundary conditions (BC's). There are several types of boundary conditions, modeling different kinds of physical constraints at the boundaries. The two most common types of BC's are *Dirichlet* and *Neumann* conditions. Given a boundary $\xi_i = 0$ or L, the Dirichlet condition prescribes a certain value for the scalar field:

$$x(\xi_i, t) = \psi_i(t). \tag{2}$$

The Neumann condition on the other hand imposes a certain flux at the boundary, i.e.:

$$x_\xi(\xi, t)_{|\xi=\xi_i} = \psi_i(t). \tag{3}$$

In more general cases, the diffusion equation can present also the lower order derivative terms $x_\xi(\xi,t)$, $x(\xi,t)$ and a source $u(\xi,t)$, that is:

$$x_t(\xi,t) = \theta_1 x_{\xi\xi}(\xi,t) + \theta_2 x_\xi(\xi,t) + \theta_3 x(\xi,t) + u(\xi,t). \tag{4}$$

The parameter θ_2 is sometime called *velocity*, while θ_3 is referred as *dissipation* or *dispersion*. In some scenarios the parameters $\{\theta_i\}$ can be also space and time varying. Under particular hypothesis on the boundaries and on the parameters (they must be constant), the diffusion equation can be solved analytically. Otherwise numerical methods are needed.

In order to solve the diffusion equation by means of numerical methods, the derivatives must be approximated with finite difference as:

$$x_\xi(\xi,t)\big|_{\xi=ih} \approx \frac{x_{i+1}(t) - x_i(t)}{h} \tag{5}$$

and

$$x_{\xi\xi}(\xi,t)\big|_{\xi=ih} \approx \frac{x_{i+1}(t) + x_{i-1}(t) - 2x_i(t)}{h^2}, \tag{6}$$

where h is the spatial sampling period. The application of the above finite differences to equation (1) at a generic point of discretization $(\xi,t) = (ih, kt_s)$ gives:

$$\begin{aligned} x_i(k+1) &= \\ &= \frac{Dt_s}{h^2}[x_{i-1}(k) - 2x_i(k) + x_{i+1}(k)] + x_i(k), \end{aligned} \tag{7}$$

where t_s is the sampling time and $x_i(k) := x(ih, kt_s)$. It can be shown that the sampling time t_s must obey the following inequality for the discrete approach to converge [2]:

$$t_s < \frac{h^2}{2D}. \tag{8}$$

Since the size of the sampling space can be also constrained by the profile of the scalar field, the inequality in (8) can pose a severe restriction on the sampling time.

IV. Problem Formulation

We consider a finite set of sensor nodes $S = \{s_i\}$ deployed over the one-dimensional sensor field $[0, L]$. The nodes are measuring discrete-time samples of the scalar field $x(\xi,t)$, continuous in time, for $t > 0$ and space for $0 < \xi < L$. It is assumed that the evolution in time and space of $x(.)$ can be modeled by the following sourceless diffusion equation:

$$x_t(\xi,t) = \theta_1 x_{\xi\xi}(\xi,t) + \theta_2 x_\xi(\xi,t) + \theta_3 x(\xi,t), \tag{9}$$

having IC $x_0(\xi)$. The boundaries, $\xi = 0$ and $\xi = L$, are subject either to Dirichlet or to Neumann conditions, not necessarily homogeneous. Notice that even if we did not include a source term in our model (9), the BC's can be used to model external sources. Hence the model we consider can reflect many real scenarios. The topic of the joint source and parameter estimation is an open problem for future research.

We assume that the node measurements are corrupted by additive zero mean white Gaussian noise, to model thermal noise from the analog/digital circuitry connected to the sensor device and also the finite precision effect of the A/D converter. Thermal noise usually dominates over the quantization effect. Furthermore the nodes $\{s_i\}$ know their absolute location $\{p_i\}$ in the field and they are partitioned into clusters $\{S_j\}$. We also assume that there is reliable communication among the nodes (i.e. no packets get lost) and that within a cluster, member nodes can talk simultaneously to the leader without interference. The task of the sensor network is to identify the parameters $\{\theta_j\}$, $j = 1, 2, 3$, at some of the nodes $\{s_i\}$.

Approaches to solve the aforementioned problem can have many applications in sensor networks. For instance, the parameters can be estimated in return to a query from the BS in order to study the physical properties of a particular environment. Moreover, the knowledge of the parameters, along with IC's and BC's, allows us to solve the PDE problem. Therefore, the estimates of the PDE coefficients could be used by the nodes the to perform in-network prediction tasks (e.g. via standard Kalman filtering). In such a scenario, only relevant events need to be notified to the base station. Alternatively the estimates could be sent to the BS (together with IC's and BC's) to supply the information necessary to reconstruct and predict the phenomenon without the need for the sensor network to send raw data continuously to the BS. Note that the latter option requires the transmission of a smaller amount of data with respect to data gathering also because the rate of variation of the parameters in time and space is slower than the one of the scalar field itself. Furthermore the estimates can be used to perform coding schemes [4].

V. Distributed Identification of Diffusion Parameters: Case of Known Boundaries

This section describes our approach to the identification of parameters of diffusion models. Here we assume that the boundary conditions are known a priori and that this

knowledge is embedded in the nodes of the sensor network. This may be the case of some controlled plant application scenarios or when the sensor region overextends the phenomenon region (here homogeneous Dirichlet BC's can be assumed). Under those assumptions a global discrete model for the physical phenomenon being monitored can be considered. The proposed discrete time-space model is presented in Subsection V-A. The algorithm to identify system parameters is derived in Subsection V-B. The extension to the case of spatially varying parameters is considered in Subsection V-C.

A. Model Discretization and Clustering

We derive a discretized model for the PDE problem described in equation (9) to be adopted by the nodes in order to identify the parameters. We assume that the information on the exact BC's is provided to the nodes (e.g. by means of some information flooding mechanisms). The nodes are partitioned into clusters, either statically or dynamically. The identification process is performed at some of the cluster heads. A cluster head receives measurements from the member nodes and process them adaptively (this is a data fusion approach). Here we do not specify the mechanisms to select the cluster heads nor to select the active clusters. For instance, some distributed election techniques for node clustering are described in [11].

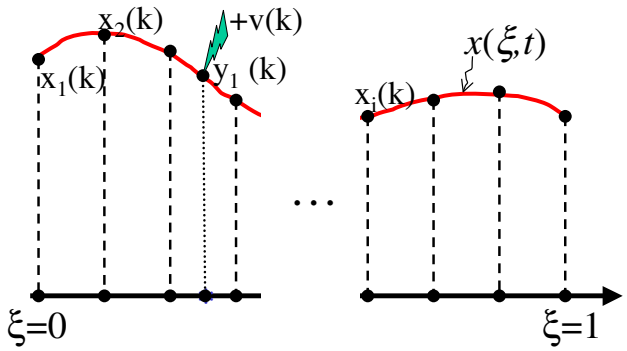

Fig. 1. State variables, $x_i(k)$, and a noisy sensor measurement $y_1(k)$ for the monodimensional scalar field $x(\xi, t)$.

The scalar field $x(\xi, t)$ is sampled in space and time at the points (ih, kt_s) where h and t_s are respectively the sampling periods in space and time. i and k are integers. The discrete samples $x_i(k)$ represent the state of the system. Defined $N := \frac{L}{h}$, there are $N + 1$ sampling points in the sensor field. We will show in the next paragraphs that the number of state variables can be less or equal to $N + 1$, depending on the type of BC's. The sampling time t_s must be selected according to the inequality constraint in (8) and therefore it is usually smaller than h.

The discretization in space and time of the PDE given in eq. (9) requires approximating spatial and time derivatives by means of finite differences, eqs. (5-6). The complete discretized model, called *lumped system*, includes the noisy sensor measurements. Figure 1 illustrates the spatial sampling of a scalar field.

For a generic point ih not adjacent to the boundaries, the expression for $x_i(k)$ is similar to the one in equation (7), but there may be more terms if lower order space derivatives are also included in the model. If the point is next to the boundaries, i.e. for $i = 1$ or $i = N$, the effect of the BC's need to be considered. We express Dirichlet boundary condition, eq. (2), by introducing the boundary value as virtual source term in the state equation. Hence, for example we define $x_1(k)$, which is next to the boundary point $\xi = 0$ where $x_0(k) = \psi_0(t)_{|t=kt_s}$, as:

$$x_1(k+1) = \theta_1 \frac{t_s}{h^2}[x_0(k) - 2x_1(k) + x_2(k)] + x_1(k),$$

$$= \theta_1 \frac{t_s}{h^2}[-2x_1(k) + x_2(k)] + x_1(k) + \theta_1 \frac{t_s}{h^2}\psi_0(kt_s). \quad (10)$$

In approximating a Neumann BC, eq. (3), we must include the boundary variable $x(\xi_i, t)$ among the state variables, because this is not constrained as for the Dirichlet BC. Then we express $x(\xi_i, t)$ with a finite difference approximation of the actual BC. For example, the BC

$$x_\xi(\xi, t)_{|\xi=0} = \psi_0(t)$$

can be approximated in the discrete space as:

$$x_\xi(\xi, t)_{|\xi=0} \approx \frac{x_1(t) - x_0(t)}{h}$$

and leads to the expressions:

$$\begin{aligned} x_0(k+1) &= \theta_1 \frac{t_s}{h^2}[x_0(k) - 2x_1(k) + x_2(k)] + x_1(k) \\ &\quad -h\psi_0(t) \quad (11) \\ x_1(k+1) &= \theta_1 \frac{t_s}{h^2}[x_0(k) - 2x_1(k) + x_2(k)] + x_1(k). \end{aligned}$$

The discretization in space and time of PDE and measurements leads to the lumped model described by the following state-space equations:

$$\mathbf{x}(k+1) = \mathbf{A}(\boldsymbol{\theta})\mathbf{x}(k) + \mathbf{B}(\boldsymbol{\theta})\mathbf{u}(k), \quad (12)$$

$$\mathbf{y}_j(k) = \mathbf{C}_j\mathbf{x}(k) + \mathbf{v}(k), \quad (13)$$

where j is the index of a cluster S_j. The state vector $\mathbf{x}(k)$ represents the uniformly sampled version of the scalar field $x(\xi, t)$, with $x_i(k) := x(ih, kt_s)$. The state matrix $\mathbf{A}(\boldsymbol{\theta})$ is a square matrix dependent on the spatial derivatives of $x(.)$ in the right hand side of eq. (9), the BC's, the sampling periods h and t_s and the set $\boldsymbol{\theta}$ of parameters to identify. If the parameters are constant over the space, $\mathbf{A}(\boldsymbol{\theta})$ is given by:

$$\mathbf{A}(\boldsymbol{\theta}) = t_s \sum_{m=1}^{3} \theta_m \mathbf{A}_i + \mathbf{I}, \quad (14)$$

where $\mathbf{I}$ is the identity matrix and the matrices $\mathbf{A}_i$ are obtained from finite difference approximation of the spatial derivatives on the right hand side of equation (9).

Equation (13) expresses the sensor measurements $\mathbf{y}_j(k)$, in a cluster S_j, as a linear interpolation, $\mathbf{C}_j\mathbf{x}(k)$, of the

state variables which is affected by the AWGN term $\mathbf{v}(k)$. Note that in general the nodes have random locations with respect to the sampling points $\{\xi = ih : i = 1, 2, ...\}$. So in most of the cases eq.(13) is a linear approximation of the relationship between the state variables and the noisy sensor measurements. In general eq. (13) should be written as:

$$\mathbf{y}_j(k) = \mathbf{g}_j(\mathbf{x}(k)) + \mathbf{v}(k), \tag{15}$$

where $\mathbf{g}_j(\mathbf{x}(k))$ is a non linear vectorial function of the state variables. Hence a polynomial interpolation may approximate $\mathbf{g}_j(.)$ better than a linear one.

Example
In this example, we derive a lumped model for the infinite dimensional system described by the diffusion equation:

$$x_t(\xi, t) = \theta x_{\xi\xi}(\xi, t),$$

where $t > 0$ and $0 < \xi < L$ with mixed Dirichlet and Neumann boundary conditions i.e.:

$$x(0, t) = \psi(t),$$

$$x_\xi(L, t) = 0.$$

There are two sensors in the field, respectively, in the locations $\xi_1 = 3L/2$ and $\xi_2 = 3L/5$, the spacial sampling rate is $h = L/5$, and there is only one cluster formed by the two sensors.

The following 5×5 state matrix $\mathbf{A}(\theta)$ can be derived in a straightforward manner through the finite differences approximation presented in Equation (6). That is,

$$\mathbf{A}(\theta) = \frac{t_s\theta}{h^2}\begin{bmatrix} -2 & 1 & 0 & 0 & 0 \\ 1 & -2 & 1 & 0 & 0 \\ 0 & 1 & -2 & 1 & 0 \\ 0 & 0 & 1 & -2 & 1 \\ 0 & 0 & 1 & -2 & 1 \end{bmatrix} + \mathbf{I},$$

$$(\mathbf{B}(\theta)\mathbf{u}(t)) = \frac{\theta}{h^2}\begin{bmatrix} \psi(t) & 0 & \dots & 0 \end{bmatrix}^T$$

and the measurement matrix $\mathbf{C}$ can be written as:

$$\mathbf{C} = \begin{bmatrix} \frac{1}{2} & \frac{1}{2} & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \end{bmatrix} \blacktriangle$$

In the above framework, some kind of information flooding mechanism id needed so that the active nodes can share the knowledge of all the model parameters: e.g. location of the boundaries and spatial sampling rate h to determine matrix $\mathbf{C}_j$. Furthermore, under the assumptions of known boundaries the nodes share the same state matrix $\mathbf{A}(\boldsymbol{\theta})$, but have different measurement equations given by (13). In an application scenario,the selection of initial parameters could be led by the BS station or a node leader in the sensor network to avoid mismatch in model definition.

B. Parameter Estimation via Kalman Filtering

The next step consists in the adaptive identification of the unknown parameters $\boldsymbol{\theta}$. For this purpose, we adopt the extended Kalman filter (EKF) [12], [10]. The key idea of the EKF approach is to treat unknown parameters as additional state variables, by defining an *augmented system*, where the state vector is defined as

$$\mathbf{z}(k) := \begin{bmatrix} \mathbf{x}(k) \\ \boldsymbol{\theta} \end{bmatrix}.$$

The augmented system is non linear, since parameters $\boldsymbol{\theta}$ are multiplied by state variables $\mathbf{x}(k)$. Here, the Kalman filter is used as the state predictor to the linearized version of the augmented system and the estimates of parameters $\boldsymbol{\theta}$ are obtained because those are treated as additional state variables in the augmented system. The augmented system can be written as

$$\begin{aligned} \mathbf{z}(k+1) &= \mathbf{f}(\mathbf{z}(k), \mathbf{u}(k)) \\ &= \begin{bmatrix} \mathbf{A}(\boldsymbol{\theta})\mathbf{x}(k) + \mathbf{B}(\boldsymbol{\theta})\mathbf{u}(k) \\ \boldsymbol{\theta} \end{bmatrix}, \qquad (16) \\ \mathbf{y}_j(k) &= \mathbf{C}_j\mathbf{x}(k) + \mathbf{v}(k). \qquad (17) \end{aligned}$$

In order to apply the Kalman filter, the Jacobians of the state and measurement equations must be derived. Those are defined as follows from equations (16) and (17):

$$\begin{aligned} \mathbf{J}(k) &:= \left.\frac{\partial \mathbf{f}(\mathbf{z}(k)\mathbf{u}(k))}{\partial \mathbf{z}}\right|_{\mathbf{z}=\hat{\mathbf{z}}}, \qquad (18) \\ \mathbf{H}_j &:= [\mathbf{C}_j \;\; \mathbf{0}]. \qquad (19) \end{aligned}$$

Then, the Kalman filter can be applied. Given the mean of the state vector, $\bar{\mathbf{z}}_0$, the initialization for the covariance matrix $\mathbf{P}$ can be found conventionally as

$$\mathbf{P}_0 = E[(\mathbf{z} - \bar{\mathbf{z}}_0)(\mathbf{z} - \bar{\mathbf{z}}_0)^T]. \tag{20}$$

Therefore, at each measurement time, the filter equations are [12]:

$$\begin{aligned} \mathbf{L}(k) &= \mathbf{P}(k/k-1)\mathbf{H}^T\left(\mathbf{H}\mathbf{P}(k/k-1)\mathbf{H}^T + \mathbf{R}\right)^{-1} \qquad (21) \\ \mathbf{P}(k/k) &= \mathbf{P}(k/k-1) - \mathbf{L}(k)\mathbf{H}\mathbf{P}(k/k-1) \qquad (22) \\ \hat{\mathbf{z}}(k/k) &= \hat{\mathbf{z}}(k/k-1) + \mathbf{L}(k)\left(\mathbf{y}(k) - \mathbf{C}_i\hat{\mathbf{x}}(k/k-1)\right) \qquad (23) \end{aligned}$$

where $\mathbf{R}$ is the covariance matrix of noise measured by sensors and the fact that $\mathbf{H}\hat{\mathbf{z}} = \mathbf{C}\hat{\mathbf{x}}$ is applied in the derivation of Equation (23). Then, the state estimate and the covariance are propagated to the next measurement time via

$$\begin{aligned} \mathbf{P}(k+1/k) &= \mathbf{J}(k)\mathbf{P}(k/k)\mathbf{J}^T(k) \qquad (24) \\ \hat{\mathbf{z}}(k+1/k) &= \mathbf{f}(\hat{\mathbf{z}}(k), u(k)). \qquad (25) \end{aligned}$$

Ljung [13] proved that the extended Kalman filter may converge to biased estimates of parameters. This is essentially due to the fact that the state estimates are performed based on the linear projection of a nonlinear system.

C. Space Varying Parameters

In some scenarios the parameters can vary with the space. This may model for example a physical medium having variable diffusivity. A parameter varying with the space ($\theta(\xi)$) will result in a set of parameters (θ_i) in the state space matrix $\mathbf{A}(\boldsymbol{\theta})$. Each parameter θ_i in the lumped system corresponds to the sample of the parameter $\theta(\xi)_{|\xi=ih}$ in the infinite dimensional system.

Example
Consider an infinite dimensional system described by the following diffusion equation. Here, the diffusion coefficient θ is space variable

$$x_t(\xi,t) = \theta(\xi)x_{\xi\xi}(\xi,t).$$

Then, the corresponding state matrix becomes

$$\mathbf{A}(\boldsymbol{\theta}) = \frac{1}{h^2}\begin{bmatrix} -2\theta_1 & \theta_1 & 0 & \dots & 0 \\ \theta_2 & -2\theta_2 & \theta_2 & 0 & \dots \\ \vdots & \vdots & \vdots & \ddots & \ddots \\ 0 & \dots & 0 & \theta_n & -2\theta_n \end{bmatrix} \blacktriangle$$

From the above example it can be seen that the general expression of the system model in equations (12) and (13) is valid even if the coefficients θ_i vary with the space.

VI. Distributed Identification: Case of Unknown Boundaries

In many realistic sensor network scenarios the a priori knowledge of the BC's may not be available. Therefore it may be necessary to sample on line the boundaries of the node region. Here we propose a model where each cluster acquires the BC's through the samples of the nodes located at its boundaries. This implies the definition of local state space models for each cluster.

Here we assume that the BC's are unknown. All the other assumptions given in Section V remain valid. The knowledge of the BC's is necessary in order to define the state space model and therefore to identify the parameters of the diffusion model. A possible approach to estimate the BC's is to have some of the nodes in charge of sampling the boundaries. We consider here a more general case where the boundaries can be defined not only globally, for the sensor field, but also locally, for a specific subregion. In order to save on communication energy and also keep a low latency, it appears reasonable to have the cluster heads receiving the sampled boundary information from nearby nodes (one hop away). In this framework, the monitored phenomenon is "sliced" in many subregions, covered by clusters and a local lumped model is defined in each cluster.

Consider a cluster of nodes $S_j \subseteq S$, and its subset of boundary nodes b_j. In one-dimension the boundary locations are indicated respectively as $\xi_l^{(j)}$ and $\xi_r^{(j)}$ (see Figure 2). The local PDE problem is defined as equation (9) but only in the space interval $\xi_l^{(j)} < \xi < \xi_r^{(j)}$. Besides, the boundary conditions are known through the noisy samples of the the nodes b_j and can be considered as Dirichlet BC's. That is the estimated boundaries

$$\hat{x}(\xi_l,t) = x(\xi_l,t) + w_l(t) \quad (26)$$
$$\hat{x}(\xi_r,t) = x(\xi_r,t) + w_r(t), \quad (27)$$

where $w_l(t)$ and $w_r(t)$ are AWGN terms. Therefore, after defining the local sampling space and time: h_j and t_j, and imposing the above BC's, the state space model for the cluster S_j can be written as:

$$\mathbf{x}_j(k+1) = \mathbf{A_j}(\boldsymbol{\theta_j})\mathbf{x}_j(k) + \mathbf{B_j}(\boldsymbol{\theta})(\mathbf{u}_j(k) + \mathbf{w}_j(k)), \quad (28)$$
$$\mathbf{y}_j(k) = \mathbf{C}_j\mathbf{x}_j(k) + \mathbf{v}(k). \quad (29)$$

where $\mathbf{w}(k)$ is a vector zero mean white Gaussian noise affecting the state of the system. If the parameters are not space varying, $\boldsymbol{\theta}_i = \boldsymbol{\theta}$.

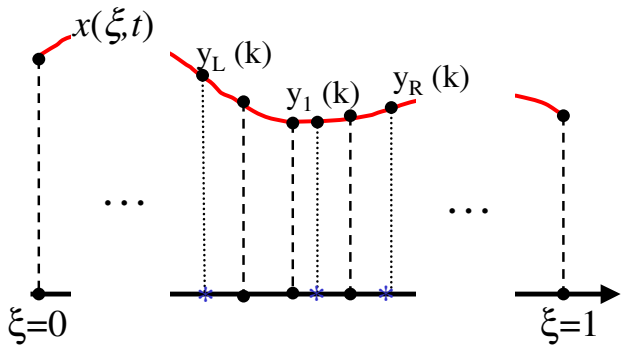

Fig. 2. Local model: The sampled BC's, $y_l(k)$ and $y_r(k)$, are sent to the cluster head.

The implication on the EKF equations of the noise affecting the state in eqs. (28) and (29) is that the update of the covariance matrix, eq. (24) in Subsection V-B now includes also a term related to the covariance matrix, $\mathbf{Q}$ of the state noise. That is:

$$\mathbf{P}(k+1/k) = \mathbf{J}(k)\mathbf{P}(k/k)\mathbf{J}^T(k) + \mathbf{B}(\boldsymbol{\theta_j})\mathbf{Q}\mathbf{B}(\boldsymbol{\theta_j})^T. \quad (30)$$

The above framework, based on local models, can be suitable for the identification of spatially varying parameters. If the variation of the parameters is slow enough we may assume it to be constant over a limited subregion. Therefore we may apply the model defined in eqs. (28) and (29) to estimate the parameters as if they were constant over a cluster. Then the global space varying parameters can be reconstructed by interpolating the estimates performed at the clusters, as it is shown in the experiments (Subsec.VIII-C).

VII. Going Toward Real Scenarios

So far, we have made some simplified assumptions, such us the monodimensionality of the sensor field and the nodes knowing their exact locations. For the particular case of

the BC's, we moved from the case of assumptions not always realistic, i.e. BC's information available at the nodes (Sec. V), to considering some nodes sampling the boundaries (Sec. VI). Here we discuss on the possible issues when relaxing some assumptions to a less ideal scenario and also on some implementation aspects of our framework.

Data gathering is one of the possible applications for this research. In a sourceless scenario, the sensor network should update the base station with the PDE parameters, the initial and boundary conditions. We note that our approach can provide IC's at an arbitrary time instant after the convergence, as it uses a joint state and parameter prediction method (the EKF). So the IC information can be aggregated to parameter estimates for an efficient delivery to the BS. On the other hand, updating the BS with boundary conditions, can be more demanding in terms of communication cost. However, this should involve only a subset of the nodes and therefore should be cheaper than performing data gathering throughout the whole network.

We mentioned that the sampling time t_s for the estimation algorithm is constrained to the inequality (8). We did not investigate yet whether t_s is much smaller than the Nyquist sampling time. If this is the case, than parameter estimation could be more energy demanding than data gathering. However, it must be pointed out that the parameter estimation may involve only a subset of the nodes and a for limited amount of time, as it stops when the algorithm converges. Therefore the cost due to an eventual higher sampling rate for the preliminary estimation of the parameters should be compensated thanks to a smaller amount of nodes involved and a shorter sampling time. Also the number of hops should be lower, as the communication stays only within the clusters. A better quantification of those trade-offs is one of the goals in our future research.

In a lot of real applications, the problem should considered at least in 2D. This implies a larger size of the matrices used in the EKF approach, but not necessarily an increase in the number of parameters to be estimated (e.g. if the medium is homogeneous). We plan to investigate methods to simplify the computation in higher than monodimensional cases.

The uncertainty regarding the node locations may bring a mismatch between the sampled data and the model adopted by the EKF. If this is above a certain degree, the correct convergence of the algorithm could be affected. Notice that there is a similar problem also related to the linear interpolation of the state variables in the measurement equation of the lumped model (12 - 13). The correct measurement equation should have a non linear term for the state variables $\mathbf{x}(k)$. This matter is also of interest for our future investigations.

VIII. Experimental Results

The goal of the experiments is to evaluate the performance of the algorithms under different conditions that may arise in a sensor network scenario. In particular, we want to understand the relationship between performance and location of the sampling point in the field (Subsec. VIII-A), the performance when the sampled scalar field has an initial large spatial bandwidth (Subsec. VIII-B) and the ability of estimating space varying parameters (Subsec. VIII-C). The results are discussed in Subsection VIII-D.

A. Locating the Sampling Point

In this experiment, we want to study how the quality of the estimate varies with respect to the noise and to the location of the nodes sampling the field. In this set up, a single node estimates the diffusivity parameter from its own noisy readings. We measure how the estimation error varies with the location of the node in the space and the level of the nodes.

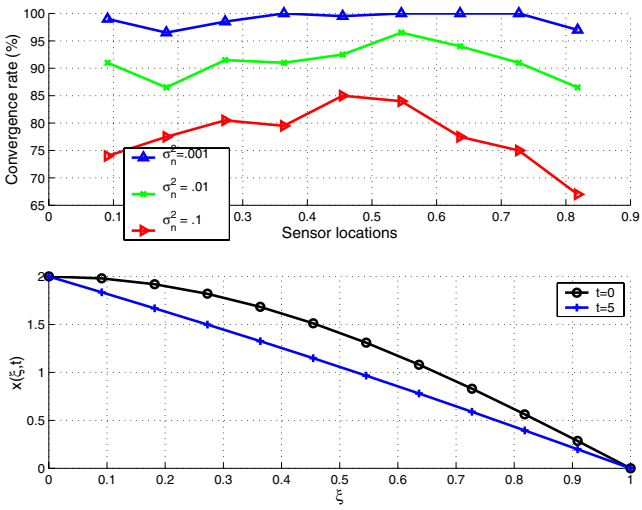

Fig. 3. Convergence rate of the algorithm versus sensor location and noise level (top). Profile of the scalar field at the initial state and at the steady state (bottom).

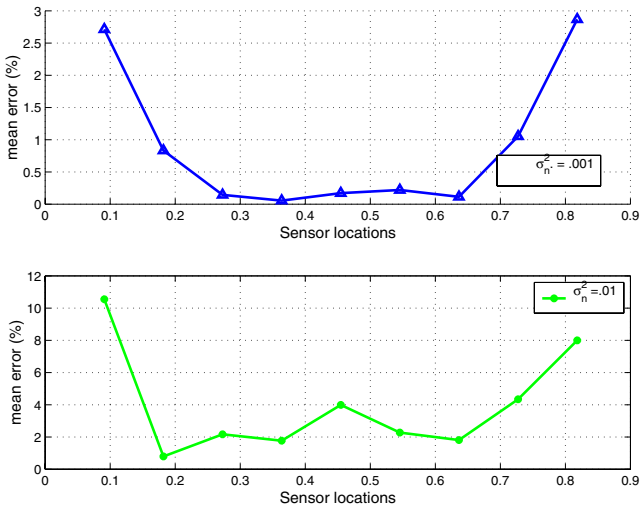

Fig. 4. Percent mean estimation error versus sensor location and noise level.

The measured scalar field is modeled by the diffusion

equation

$$x_t(\xi,t) = \frac{1}{\pi^2} x_{\xi\xi}(\xi,t),$$

defined in the interval $0 < \xi < 1$ and $t > 0$, with initial condition $x(\xi,0) = 2\cos(\frac{\pi}{2}\xi)$ and Dirichlet boundary conditions: $x(0,t) = 2$, $x(1,t) = 0$.

A single node s in position p is estimating the diffusivity parameter with sampling time $t_s = 0.004$, about 1/10th of the right hand side of inequality given by eq. (8). Note that the sampling space of the lumped model (12)-(13) is set to 1/11. We assume that the node knows the BC's. The position p of s varies uniformly on the ξ axis. Different noise levels, σ_n^2, are considered. The average percent error is measured over 200 Monte Carlo trials and for $p = k/11, k \in \{1, 2, ..., 10\}$.

We notice that the algorithm does not always converge. The divergent iterations are excluded from the computation of the mean error. In this set up, the estimates are performed only by one single node. This explains the relatively high divergence rate in the noisy cases. The performance would have been much better by processing data from multiple nodes. Besides increasing the number of sensors, more robust estimates may be obtained also through some modified versions of the EKF. The percentage of convergent trials is shown in the upper part of Figure 3 while the mean error w.r.t. the sensor location is shown in Figure 4 for two different levels of noise.

The profiles of the scalar field at the initial time and at the steady state are shown in Figure 3. It can be noticed that the performance degrades as the measurements are taken closer to any of the boundaries. and it is not simply directly related to the SNR. The SNR is monotonically decreasing when approaching the point $\xi = 1$, but the estimation performance seems depending also on the rate of temporal variability of the sample scalar field.

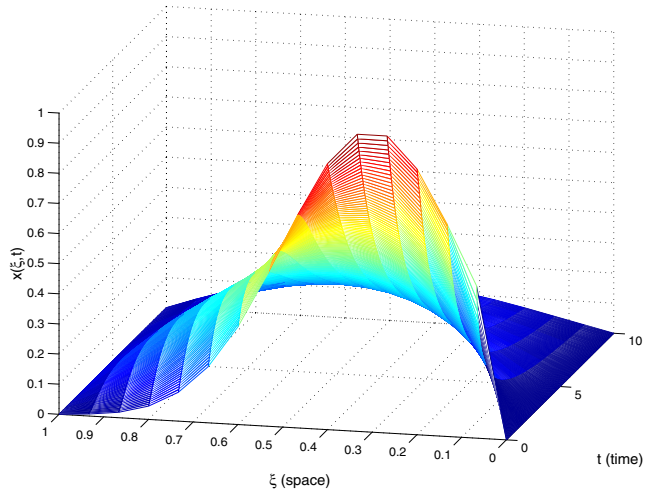

Fig. 5. Scalar field described by the parabolic equation (31). An advection phenomenon can be observed: the concentration is moving toward one of the boundaries.

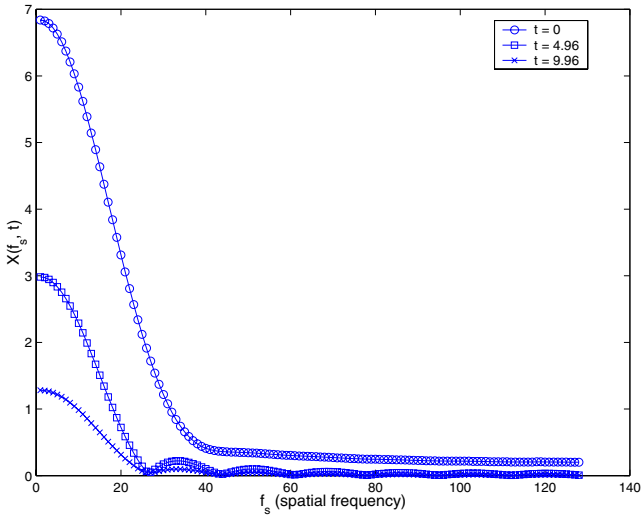

Fig. 6. Fourier transform with respect to the space at three different time instants, to show the smoothing effect.

B. Gaussian Pulse

Here we consider the parabolic equation

$$x_t(\xi,t) = \frac{1}{50} x_{\xi\xi}(\xi,t) - \frac{1}{50} x_\xi(\xi,t) - .32x(\xi,t). \qquad (31)$$

Since all the coefficients are nonzero, the equation can model a phenomenon of advection-diffusion with dissipation. In other words, besides a diffusion, there is a motion of the concentration (advection) $x(\xi,t)$, due to the term $-\frac{1}{50}x_\xi(\xi,t)$ and also a dissipation due to the term $-.32x(\xi,t)$. The initial conditions is: $x(\xi,0) = \exp(-(\xi - .3).^2/(.08))$, with homogeneous Dirichlet BC's: $x(0,t) = x(1,t) = 0$. Here $x(\xi,0)$ has a larger bandwidth w.r.t. the I.C.'s of the problem in Subsection VIII-A.

The field $x(\xi,t)$ is shown in Figure 5, while the smoothing property (Sec. III) of the diffusion equation is highlighted through Figure 6: the spatial bandwidth of the field is getting narrower with the time. A cluster of 5 nodes centrally located is used. After 200 Monte Carlo iterations, the percent mean error plus standard deviation in the estimation of the three parameters are respectively: $1.70 + 3.41$, $1.90 + 8.01$ and $9.38 + 21.7$. The estimates of the PDE parameters are displayed in Figures 8 and 9.

C. Space Varying Parameters

In this section we address the problem of estimating space varying parameters under unknown B.C.'s. We consider the PDE system

$$x(\xi,t) = (\theta(\xi)x_\xi(\xi,t))_\xi \qquad (32)$$

in the interval $0 < \xi < 1$ and $t > 0$, with initial conditions: $x(\xi,0) = 2\cos(\frac{\pi}{2}\xi)$ and Dirichlet boundary conditions $x(0,t) = 2$, $x_\xi(1,t) = 0$. This time the diffusion coefficient is variable with the space and is defined as:

$$\theta(\xi) = \theta_0 + \theta_1 f(\xi) = .1 - .2(\xi - .5)^2.$$

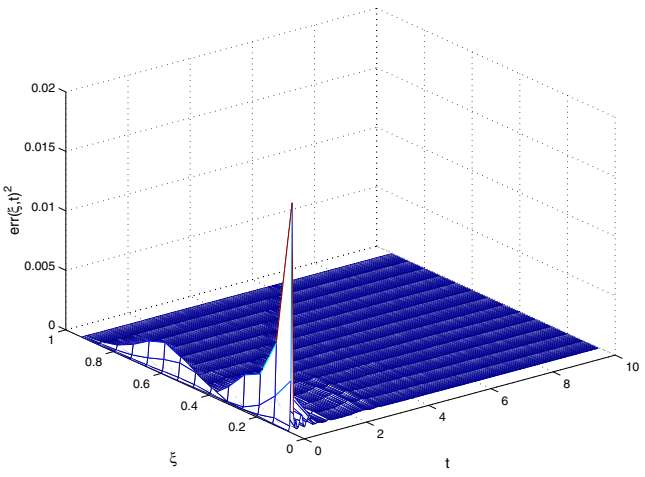

Fig. 7. Square error in the estimation of the scalar field.

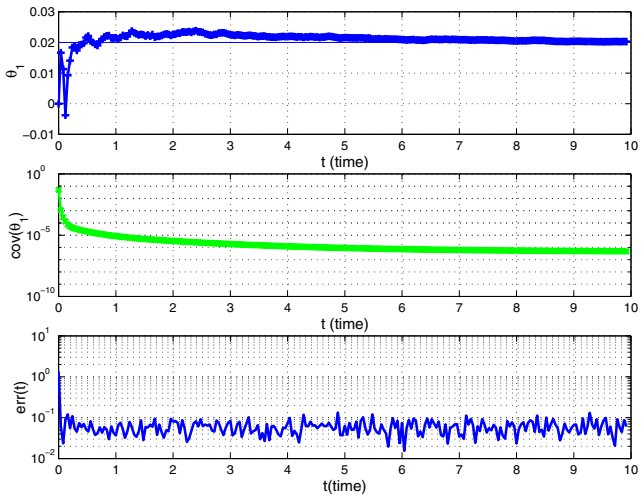

Fig. 8. Estimate of diffusion coefficient (top), related covariance (middle) and prediction error (bottom) related to the parabolic problem in eq. (31).

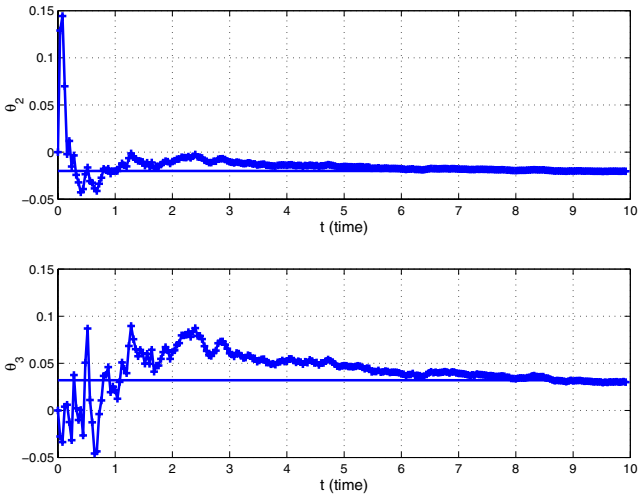

Fig. 9. Estimates of coefficients θ_2 (top) and θ_3 (bottom).

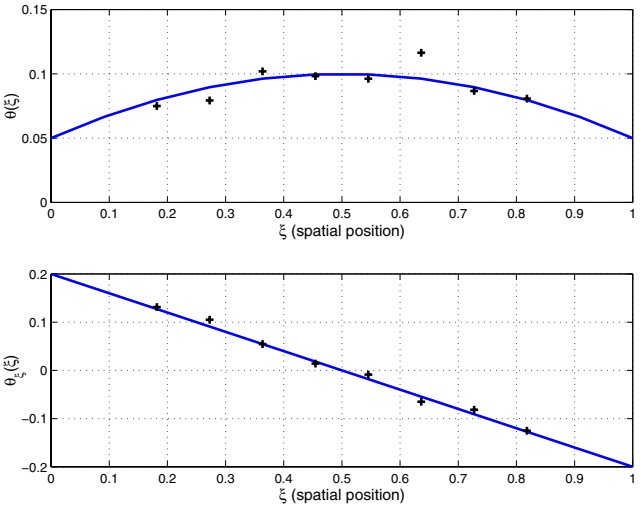

Fig. 10. Comparison of the local estimates of the parameters with the true continuous values.

Thus eq. (32) can be rewritten as:

$$x(\xi,t) = \theta(\xi)x_{\xi\xi}(\xi,t) + \theta_{\xi}(\xi)x_{\xi}(\xi,t). \qquad (33)$$

To identify the coefficients, we partition the sensor field into clusters and we assume the parameters to be constant over that cluster. The boundary conditions (time varying) are acquired by the boundary nodes. The estimates of the parameters are compared to the true coefficients in Figure 10, to show their accuracy.

D. Discussion

Several factors affect the performance of non centralized estimates of parameters of diffusion models. As expected, the SNR, the sampling step of the estimator, the number of sensors are among those factor. It must be also pointed out that the location of the sensors in the region plays an important role. This is not only because the signal level, and therefore the SNR is different from point to point, but also because the variability of the field over some areas is higher and therefore allows better identification of the parameters.

We also verified that multiple local models with constant parameters can be used to identify (and hence to approximate) space varying parameters. This allows to have a limited number of parameters to be estimated at each cluster.

It can be noticed how in general the convergence of the algorithm is quite fast. Hence only a relatively small number of iterations is needed to estimate the parameters (at least as long as those are constant over the time).

IX. Conclusion

We have focused on the topic of monitoring diffuse phenomena via autonomous sensor networks. In our work, diffusion partial differential equations (PDE's) were adopted to model the time and space correlations of the physical phenomena observed by the sensor field. In this scenario, the sensor network can send only the model parameters, the initial and boundary conditions to the base station (BS).

This information suffices to predict the evolution of the phenomenon without the need of continuous updates of raw data from the sensor field to the BS. Our main focus here has been on the distributed identification of parameters of PDE's under various types of possibly unknown boundary conditions. To this end, we presented a method based on extended Kalman filtering (EKF). We have also addressed the problem of the identification of spatial varying parameters. Simulation results to study the performance of this scheme have been shown and discussed.

Future work will try to study more in depth how different factors may affect the convergence and the performance of the algorithm. We will also focus on the analysis of the communication cost factors and on the methods to treat possibly unknown sources in the sensor field.

X. Acknowledgements

This work is supported in part by NSF grants number 0347621 and 0325875.

We would like to thank our anonymous reviewers for their feedback. Among the other things, that helped us in refining the directions for our future research.

References

[1] W. Rabiner Heinzelman, A. Chandrakasan, and H. Balakrishnan, "Energy efficient communication protocol for wireless sensor networks," in *Proc. International Conference on System Sciences*, January 2000, vol. 33.

[2] R. Ghez, *Diffusion Phenomena*, Kluwer Academic/Plenum Publishers, 2nd edition, 2001.

[3] L. A. Rossi, B. Krishnamachari, and C.-C. Jay Kuo, "Monitoring of diffusion processes with pde models in wireless sensor networks," in *Defense and Security Symposium 2004*, April 12-16 2004.

[4] B. Beferull-Lozano, R. L. Konsbruck, and M. Vetterli, "Rate-distortion problem for physics based distributed sensing," in *Information Processing in Sensor Networks (IPSN)*, April 26-27 2004.

[5] Y. Orlov and J. Bentsman, "Adaptive distributed parameter systems identification with enforceable identificability conditions and reduced-order spatial differentiation," *IEEE Trans. on Automatic Control*, vol. 45, no. 2, pp. 36–50, February 2000.

[6] V. Isakov and S. Kindermann, "Identification of the diffusion coefficient in a one-dimensional parabolic equation," *Inverse Problems*, vol. 16, no. 2000, pp. 665–680, 2000.

[7] S. Xiong and Y. Li, "Evolutionary algorithm for parameter identification inverse problem in parabolic systems," in 4^{th} *World Congress on Intelligent Control and Automation*, June 2002, pp. 1817–1821.

[8] J. Baumeister, W. Scondo, M.A. Demetrioiu, and I.G. Rosen, "On-line parameter estimation for infinite dimensional dynamical systems," *SIAM J. Control Optimization*, vol. 35, no. 2, pp. 678–713, March 1997.

[9] R.F. Curtain and H.J. Zwart, *An Intoduction to Infinite-Dimensional Linear Systems Theory*, Springer Verlag, 1995.

[10] J. Yin, V.L. Syrmos, and D.Y.Y Yun, "System identification using nonlinear filtering methods with applications to medical imaging," in *39th IEEE Conference on Decision and Control*, December 2000, vol. 4, pp. 3313–3318.

[11] G. Tell, *Introduction to Distributed Algorithms*, Cambridge University Press, 1994.

[12] J. M. Mendel, *Lessons in Estimation Theory for Signal Processing, Communications and Control*, Prentice Hall, 1995.

[13] L. Ljung, "Asymptotic behavior of the extended kalman filter as a parameter estimator for linear systems," *IEEE Trans. on Automatic Control*, vol. AC-24, no. 1, pp. 36–50, February 1979.